

(12) NACH DEM VERTRAG ÜBER DIE INTERNATIONALE ZUSAMMENARBEIT AUF DEM GEBIET DES
PATENTWESENS (PCT) VERÖFFENTLICHTE INTERNATIONALE ANMELDUNG

(19) Weltorganisation für geistiges Eigentum
Internationales Büro



(43) Internationales Veröffentlichungsdatum
16. Mai 2002 (16.05.2002)

PCT

(10) Internationale Veröffentlichungsnummer
WO 02/38803 A2

- (51) Internationale Patentklassifikation⁷: **C12Q 1/68** (74) **Anwalt:** SCHÜSSLER, Andrea; Truderinger Strasse 246, 81825 München (DE).
- (21) Internationales Aktenzeichen: PCT/DE01/04229
- (22) Internationales Anmeldedatum:
8. November 2001 (08.11.2001)
- (25) Einreichungssprache: Deutsch
- (26) Veröffentlichungssprache: Deutsch
- (30) Angaben zur Priorität:
100 55 285.4 8. November 2000 (08.11.2000) DE
- (71) **Anmelder** (für alle Bestimmungsstaaten mit Ausnahme von US): **DEUTSCHES KREBSFORSCHUNGSZENTRUM STIFTUNG DES ÖFFENTLICHEN RECHTS** [DE/DE]; Im Neuenheimer Feld 280, 69120 Heidelberg (DE).
- (72) **Erfinder; und**
- (75) **Erfinder/Anmelder** (nur für US): **EICHMÜLLER, Stefan** [DE/DE]; Strassburger Ring 15, 68535 Edingen-Neckarhausen (DE). **SCHADENDORF, Dirk** [DE/DE]; Weberstrasse 3, 68165 Mannheim (DE). **USENER, Dirk** [DE/DE]; Elberstrasse 54, 55122 Mainz (DE).
- (81) **Bestimmungsstaaten (national):** AE, AG, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, BZ, CA, CH, CN, CO, CR, CU, CZ, DK, DM, DZ, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, MZ, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.
- (84) **Bestimmungsstaaten (regional):** ARIPO-Patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), eurasisches Patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), europäisches Patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE, TR), OAPI-Patent (BF, BJ, CF, CG, CI, CM, GA, GN, GQ, GW, ML, MR, NE, SN, TD, TG).
- Veröffentlicht:**
— ohne internationalen Recherchenbericht und erneut zu veröffentlichen nach Erhalt des Berichts
- Zur Erklärung der Zweibuchstaben-Codes und der anderen Abkürzungen wird auf die Erklärungen ("Guidance Notes on Codes and Abbreviations") am Anfang jeder regulären Ausgabe der PCT-Gazette verwiesen.*

(54) **Title:** NOVEL MARKER FOR THE DIAGNOSIS AND THERAPY OF TUMOURS

(54) **Bezeichnung:** NEUE MARKER FÜR DIE DIAGNOSE UND THERAPIE VON TUMOREN

(57) **Abstract:** The invention relates to a novel marker for tumours, preferably CTCL. The invention further relates to the application of the above for the diagnosis and therapy of tumour-related diseases, preferably CTCL.

(57) **Zusammenfassung:** Die vorliegende Erfindung betrifft neue Marker für Tumoren, vorzugsweise CTCL. Die vorliegende Erfindung betrifft ferner deren Einsatz für die Diagnose bzw. Therapie von Tumorerkrankungen, vorzugsweise CTCL.

WO 02/38803 A2

Best Available Copy

Neue Marker für die Diagnose und Therapie von Tumoren

Die vorliegende Erfindung betrifft den Einsatz neuer Marker für die Diagnose bzw. Therapie von Tumorerkrankungen, vorzugsweise CTCL.

Kutane T-Zell-Lymphome (CTCL) stellen eine heterogene Gruppe von Erkrankungen dar, bei denen CD4-T-Zellen als der maligne Zelltyp vorherrschen. In den meisten Fällen wird der mono- oder zumindest oligoklonale Ursprung der malignen Zellen anhand T-Zellrezeptor-Rearrangements dokumentiert. Neben verschiedenen weiteren Untertypen stellen Mycosis fungoides und das Sézary-Syndrom (SS) die häufigsten Formen von CTCL dar. Bei beiden Erkrankungen handelt es sich um monoklonale T-Helfer-Memory-Lymphome, die durch kutane Plaques, Tumore oder Erythrodermie gekennzeichnet sind, wobei SS zusätzlich durch eine generalisierte Lymphadenopathie und die Gegenwart von neoplastischen T-Zellen im peripheren Blut charakterisiert ist.

Zu den therapeutischen Ansätzen zählen die Stadium-abhängige Auswahl von PUVA (Psoralen und UV-A), Retinoiden, Interferon α -2a in Kombination mit Acitretin oder PUVA, verschiedene Immunmodulatoren, Elektronenbestrahlung oder extrakorporale Photopherese. Diese Verfahren sind in frühen Krankheitsstadien erfolgreich, nicht jedoch bei den aggressiven späteren Stadien. Zu möglicherweise sinnvollen zukünftigen Therapien für CTCL zählen immunologische Therapien, z.B. Vakzinierung mit Peptiden oder Peptid-beladenen dendritischen Zellen so wie dies bereits auch zur Behandlung von Melanomen verwendet wurde.

Die Gegenwart und Aktivität von CD8⁺-Zellen bei CTCL wurde mit der Prognose korreliert. Es konnte gezeigt werden, daß CD8⁺-reaktive Infiltrate CTCL-spezifisch und lytisch sind. Somit mögen zwar Immuntherapien ein vielversprechendes Konzept zur Behandlung CTCL darstellen, eine Voraussetzung für eine solche Strategie ist allerdings die Identifizierung Tumor-spezifischer Antigene. In diesem Zusammenhang wurde der T-Zellrezeptor selbst als ein Antigen vorgeschlagen (ähnlich den Idiotyp-Immunglobulinen als Ziel für B-Zell-spezifische T-Zellen). In

beiden Fällen ist man jedoch mit dem Nachteil konfrontiert, daß der Antigen-T-Zellrezeptor für jeden einzelnen Patienten identifiziert werden müßte. Zusammengefaßt muß allerdings festgestellt werden, daß für Tumorarten wie CTCL zur Zeit keine tumorassoziierten Antigene bekannt sind, somit die Möglichkeiten der spezifischen Diagnose bzw. Therapie stark eingeschränkt sind.

Somit liegt der vorliegenden Erfindung im wesentlichen das technische Problem zugrunde, Marker (Gene bzw. deren Produkte) zu identifizieren und bereitzustellen, die mit Tumoren, insbesondere CTCL, in Zusammenhang stehen und gegebenenfalls von diagnostischem und/oder im Rahmen einer Vakzinierungstherapie von therapeutischem Nutzen sind.

Die Lösung dieses technischen Problems wurde durch die Bereitstellung der in den Patentansprüchen gekennzeichneten Ausführungsformen erzielt.

Es wurden überraschenderweise eine Reihe von Genen gefunden, deren Expression mit CTCL korreliert ist. In den zu der vorliegenden Erfindung führenden Experimenten wurden CTCL-spezifische Antigene durch Screenen einer Testis-cDNA-Bank bzw. einer cDNA-Bank, die aus Tumor-RNA verschiedener kutaner Lymphome hergestellt worden war, mit Seren von Tumorpatienten identifiziert. Etwa 3×10^6 Rekombinante wurden mit Seren von Patienten mit Sézary-Syndrom bzw. Mycosis fungoides gescreent. Die Ergebnisse zeigen, daß Tumor-Antigene von CTCL-Tumoren mit von Tumorpatienten stammenden Antikörpern identifiziert werden können. Es konnten positive Klone identifiziert werden, die zu 19 unterschiedlichen Genen/ORFs gehörten, wozu auch fünf bisher nicht bekannte Sequenzen zählen. Alle gefundenen Tumor-Antigene sind spezifisch, d.h. daß nur Tumorpatienten, jedoch keine gesunden Personen gegen diese gerichtete Antikörper bilden, obwohl 13 von diesen Tumor-Antigenen in mindestens 21% der getesteten Kontrollgewebe exprimiert werden. Es wurde außerdem ein tumorspezifisches Antigen gefunden, das nur in Testis und Tumorgewebe exprimiert wird. Dabei handelt es sich um se2-1, das in einem CTCL-Tumor gefunden wurde. Dieses Gen zeigt

Ähnlichkeiten zu SCP-1, ein mit der Mitose in Zusammenhang stehendes Protein. Vier Seren von CTCL-Patienten reagierten mit verschiedenen SCP-1-ähnlichen Klonen. Somit konnten mittels den zur der vorliegenden Erfindung führenden Experimenten zum ersten Mal CTCL-assoziierte Antigene (unabhängig vom T-Zellrezeptor) identifiziert werden, die somit wertvolle Tumormarker darstellen. Die Identifizierung solcher Antigene ist von Interesse, da die kodierten Proteine und von diesen abgeleitete Peptide als Zielstrukturen, z.B. für zytotoxische Zellen, dienen und als Antigene zur Produktion von diagnostischen oder therapeutischen Antikörpern verwendet werden können. Zur Tumorthherapie können die von den erfindungsgemäßen Nukleinsäuren kodierten Peptide bzw. Fragmente davon entweder direkt appliziert werden oder auf Antigen-präsentierende Zellen geladen werden. Auch können die Antigene darstellenden Peptide mit Hilfe von Vektoren in verschiedenen Zellen (z. B. Dendritischen Zellen als Antigen-präsentierende Zellen) exprimiert werden. Weiter bilden die identifizierten Nukleinsäuren die Grundlage zur Entwicklung diagnostischer Tests, um zukünftig eine zuverlässigere und frühzeitige Diagnose bei Betroffenen zu gewährleisten. Darüber hinaus werden funktionelle Analysen der Proteine zweifellos zum Verständnis der Tumorentwicklung beitragen. Die erfindungsgemäßen Nukleinsäuren sind somit als Kandidatengene für Untersuchungen der Pathomechanismen anzusehen, die verschiedenen Tumorerkrankungen, wie z.B. CTCL, zugrunde liegen.

Gegenstand der vorliegenden Erfindung ist somit eine diagnostische Zusammensetzung, die mindestens eine Nukleinsäuresequenz enthält, deren veränderte Expression mit einer Tumor-Erkrankung assoziiert ist, wobei die Nukleinsäuresequenz se2-5 (Fig.1), se20-10 (Fig.2), se57-1 (Fig.3), se70-2 (Fig.4), Lg1-2 (Fig. 5), se1-1 (Fig. 6), se2-1 (Fig.7), se2-2 (Fig.8), se14-3 (Fig.9), se20-4 (Fig.10), se20-7 (Fig.11), se20-9 (Fig.12), se33-1 (Fig.13), se37-2 (Fig.14), se89-1 (Fig.15), L14-2 (Fig.16), L15-7 (Fig.17), Li9-1 (Fig.18), Li9-4 (Fig.19), Lii5-2 (Fig. 20), Lii10-6 (Fig. 21), Liii4-5 (Fig. 22) oder GBP-TA (Fig. 23) umfaßt.

Gegenstand der vorliegenden Erfindung ist ferner ein Arzneimittel, das eine Nukleinsäuresequenz enthält, deren veränderte Expression mit einer Tumor-Erkrankung assoziiert ist, wobei die Nukleinsäuresequenz se20-10 (Fig.2), se57-1 (Fig.3), Lg1-2 (Fig. 5), se1-1 (Fig. 6), se2-1 (Fig.7), se2-2 (Fig.8), se14-3 (Fig.9), se20-4 (Fig.10), se20-7 (Fig.11), se20-9 (Fig.12), se33-1 (Fig.13), se37-2 (Fig.14), L14-2 (Fig.16), L15-7 (Fig.17), Li9-1 (Fig.18), Li9-4 (Fig.19), Lii5-2 (Fig. 20), Lii10-6 (Fig. 21), Liii4-5 (Fig. 22) oder GBP-TA (Fig. 23) umfaßt.

Die vorliegende Erfindung betrifft ferner eine vorstehend definierte diagnostische Zusammensetzung oder ein Arzneimittel, wobei die Nukleinsäuresequenz, deren veränderte Expression mit einer malignen Tumor-Erkrankung in Zusammenhang steht, eine Nukleinsäuresequenz umfaßt,

- (a) die sich von einer vorstehenden, in den Figuren 1 bis 23 dargestellten Nukleinsäuresequenz in der Codonsequenz aufgrund der Degeneration des genetischen Codes unterscheidet;
- (b) die mit einer Nukleinsäuresequenz nach einer der Figuren 1 bis 23 oder nach (a) hybridisiert; oder
- (c) die ein Fragment, eine allelische Variante oder eine andere Variante einer der vorstehend definierten Nukleinsäuresequenz ist.

Der Begriff "hybridisierende Nukleinsäuresequenz" weist auf eine Nukleinsäuresequenz hin, die unter üblichen Bedingungen, insbesondere bei 20°C unter dem Schmelzpunkt der Nukleinsäure, mit einer in den Figuren gezeigten Nukleinsäuresequenz hybridisiert. Der in vorliegenden Erfindung verwendete Begriff "hybridisieren" bezieht sich auf konventionelle Hybridisierungsbedingungen, vorzugsweise auf Hybridisierungsbedingungen, bei denen als Lösung 5xSSPE, 1% SDS, 1xDenhardts-Lösung verwendet wird und/oder die Hybridisierungstemperatur zwischen 50°C und 70°C, vorzugsweise bei 65°C liegen. Nach der Hybridisierung wird vorzugsweise zuerst mit 2xSSC, 1% SDS und danach mit 0,2xSSC bei Temperaturen zwischen 50°C und 70°C, vorzugsweise bei 65°C gewaschen (zur Definition von SSPE, SSC und Denhardts-Lösung siehe Sambrook et

al., Molecular Cloning: A Laboratory Manual, 2. Ausgabe, Cold Spring Harbor Laboratory Press, Cold Spring Harbor NY (1989)). Besonders bevorzugt sind stringente Hybridisierungsbedingungen, wie sie beispielsweise in Sambrook et al., supra, beschrieben sind.

Die in der vorliegenden Erfindung verwendeten Begriffe "andere Variante" oder "Fragment" umfassen Nukleinsäuresequenzen, die sich gegenüber den in den Figuren angegebenen Sequenzen oder vorstehenden hybridisierenden Sequenzen durch Deletion(en), Insertion(en), Austausch(e) und/oder andere im Stand der Technik bekannte Modifikationen unterscheiden bzw. ein Fragment der ursprünglichen Nukleinsäuresequenz umfassen, wobei das durch diese Nukleinsäuresequenzen kodierte Protein noch eine oder mehrere der vorstehend bzw. in den Beispielen beschriebenen Eigenschaften aufweist. Dazu zählen auch Allelvarianten. Die Varianten weisen eine Homologie zu den beanspruchten Sequenzen von mind. 70%, von mindestens 80%, bevorzugt mindestens 90%, ganz bevorzugt mindestens 95%, 96%, 97%, 98% oder 99% auf. Verfahren zur Erzeugung der vorstehenden Änderungen in der Nukleinsäuresequenz sind dem Fachmann bekannt und in Standardwerken der Molekularbiologie beschrieben, beispielsweise in Sambrook et al., supra. Der Fachmann ist auch in der Lage, zu bestimmen, ob ein von einer so veränderten Nukleinsäuresequenz kodierte Protein noch über die gewünschten biologischen Eigenschaften verfügt. Vor allem in Zusammenhang mit diagnostischen Anwendungen bzw. Zusammensetzungen, in denen eine der vorstehenden Nukleinsäuresequenzen verwendet wird, bezieht sich der Begriff "Fragment" auf ein Fragment, das eine Länge von mindestens 12, vorzugsweise mindestens 20 und noch mehr bevorzugt mindestens 25 Nukleotiden aufweist.

In einer bevorzugten Ausführungsform ist das vorstehend definierte Nukleinsäuremolekül eine cDNA.

In einer weiteren bevorzugten Ausführungsform ist die Nukleinsäuresequenz eine genomische DNA, die vorzugsweise von einem Säuger, beispielsweise einem Menschen stammt. Auf Nukleinsäurehybridisierung basierende Screening-Verfahren erlauben die

Isolierung der erfindungsgemäßen genomischen DNA-Moleküle aus jedem Organismus bzw. abgeleiteten genomischen DNA-Banken, wobei Sonden verwendet werden, die die in den Figuren angegebene Nukleinsäuresequenz oder ein Fragment davon enthalten.

Die Nukleinsäuresequenzen können auch in einen Vektor bzw. Expressionsvektor inseriert werden. Beispiele solcher Vektoren sind dem Fachmann bekannt. Im Falle eines Expressionsvektors für *E. coli* sind dies z. B. pGEMEX, pUC-Derivate (z.B. pUC8), pBR322, pBlueScript, pGEX-2T, pET3b und pQE-8. Für die Expression in Hefe sind z.B. pY100 und Ycpad1 zu nennen, während für die Expression in tierischen Zellen z.B. pKCR, pEFBOS, cDM8 und pCEV4, anzugeben sind. Für die Expression in speziell Insektenzellen eignet sich besonders der Baculovirus-Expressionsvektor pAcSGHisNT-A. In einer bevorzugten Ausführungsform ist die Nukleinsäuresequenz im Vektor mit regulatorischen Elementen funktionell verknüpft, die dessen Expression in prokaryotischen oder eukaryötischen Wirtszellen erlauben. Solche Vektoren enthalten neben den regulatorischen Elementen, beispielsweise einem Promotor, typischerweise einen Replikationsursprung und spezifische Gene, die die phänotypische Selektion einer transformierten Wirtszelle erlauben. Zu den regulatorischen Elementen für die Expression in Prokaryonten, beispielsweise *E.coli*, zählen der lac-, trp-Promotor oder T7-Promotor, und für die Expression in Eukaryonten der AOX1- oder GAL1-Promotor in Hefe, und der CMV-, SV40-, RVS-40-Promotor, CMV- oder SV40-Enhancer für die Expression in tierischen Zellen. Weitere Beispiele für geeignete Promotoren sind der Metallothionein I- und der Polyhedrin-Promotor.

Zu geeigneten Vektoren zählen insbesondere auch auf T7 basierende Expressionsvektoren für die Expression in Bakterien (Rosenberg et al., *Gene* 56(1987), 125) oder pMSXND für die Expression in Säugerzellen (Lee und Nathans, *J.Biol.Chem.* 263(1988), 3521).

Allgemeine auf dem Fachgebiet bekannte Verfahren können zur Konstruktion von Vektoren oder Plasmiden, die die vorstehenden Nukleinsäuresequenzen und geeignete Kontrollsequenzen enthalten,

verwendet werden. Zu diesen Verfahren zählen beispielsweise in vitro-Rekombinationstechniken, synthetische Verfahren, sowie in vivo-Rekombinationsverfahren, wie sie beispielsweise in Sambrook et al., supra, beschrieben sind.

Wirtsorganismen können mit den vorstehend beschriebenen Vektoren transformiert werden. Zu diesen Transformanten zählen Bakterien, Hefe, Insekten- und weitere Tierzellen, vorzugsweise Säugerzellen. Bevorzugt sind die E. coli Stämme HB101, DH1, x1776, JM101, JM109, BL21, XL1Blue und SG 13009, der Hefestamm Saccharomyces cerevisiae und die tierischen Zellen L, 3T3, FM3A, CHO, COS, Vero, HeLa sowie die Insektenzellen sf9. Verfahren zur Transformation dieser Wirtszellen, zur phänotypischen Selektion von Transformanten und zur Expression der vorstehenden Nukleinsäuresequenzen unter Verwendung der vorstehend beschriebenen Vektoren sind auf dem Fachgebiet bekannt.

Die vorstehend erwähnten Nukleinsäuren eignen sich besonders als Antigen-kodierende Struktur für therapeutische Zwecke. Ziel soll es dabei sein, das Immunsystem zu stimulieren, Tumorzellen zu eliminieren, die über eine Nukleinsäure identifiziert werden. Dabei gibt es verschiedene Wege, z. B. Injektion der nackten DNA in den Patienten. Hierzu wird ein Plasmid mit einem sehr aktiven Promotor und mindestens einer erfindungsgemäßen Nukleinsäure, die insbesondere se20-10, se57-1, Lg1-2, se1-1, se2-1, se2-2, se14-3, se20-7, se20-9, se33-1, se37-2, L14-2, L15-7, Li9-1, Li9-4, Lii5-2, Lii10-6, Liii4-5 oder GBP-TA umfaßt, beispielsweise in den Muskel oder intradermal injiziert.

Desweiteren kann die Nukleinsäuresequenz nicht nur zu Zwecken der rekombinanten Herstellung in einen Vektor inseriert sein, sondern auch um die DNA mit Hilfe von Vektoren in Patienten zu injizieren, wo diese ein Antigen für therapeutische Zwecke kodiert. Das Ziel ist, daß die Zellen das Plasmid aufnehmen, Antigene produzieren, über HLA-Moleküle einzelne Peptide präsentieren und somit eine zytotoxische T-Zell-Immunantwort hervorrufen. Diese soll dann zur Abwehr von Tumorzellen führen. Allgemein ist diese Verfahrensweise in Conry et al., Clinical Cancer Research 4, S. 2903-2912 (1998) beschrieben. Einen

alternativen Weg stellt die "Gene-Gun"-Methode dar, die in Fynan et al., Proc. Natl. Acad. Sci. USA 90, S. 11478-11482 (1993) beschreiben ist. Hierbei wird die Nukleinsäure mit Hilfe eines Vektors in vivo Antigen-präsentierende Zellen (APCs), z.B. Dendritische Zellen, zur HLA-Präsentation des kodierten Proteins gebracht. Dabei kann der die erfindungsgemäße Nukleinsäure enthaltende Vektor über verschiedene Wege injiziert werden:

- a) Lipid- bzw. Liposomen-verpackte DNA oder RNA, z.B. allgemein beschrieben von Nabel et al., Proc. Natl. Acad. Sci. USA 93, S. 15388-15393 (1996).
- b) Mit einem Bakterium als Transportvehikel für den Expressionsvektor. Geeignete Bakterien sind beispielsweise (attenuierte) Listerien [z.B. *Listeria monocytogenes*], Salmonellen-Stämme [z.B. *Salmonella* spp.]. Allgemein ist diese Technik von Medina et al., Eur. J. Immunol. 29, S. 693-699 (1999) sowie Guzman et al., Eur. J. Immunol. 28, S. 1807-1814 (1998) beschrieben worden. Desweiteren wird auf WO 96/14087; Weiskirch et al., Immunological Reviews 158, S. 159-169 (1997) und US-A-5,830,702 verwiesen.
- c) Mittels Gene-Gun (Williams et al., Proc. Natl. Acad. Sci. USA 88, S. 2726-2730, 1991).

In einer bevorzugten Ausführungsform ist der die vorstehenden Nukleinsäuresequenzen enthaltene Vektor ein Virus, beispielsweise ein Adenovirus, Vaccinia-Virus oder ein AAV-Virus, der bei einer Gentherapie von Nutzen ist. Besonders bevorzugt sind Retroviren. Beispiele für geeignete Retroviren sind MoMuLV, HaMuSV, MuMTV, RSV oder GaLV. Weiter eignen sich vorgenannte Viren sowie Fowlpox-Virus, Canarypox-Virus, Influenza-Virus oder Sindbis-Virus auch als Basis einer Vakzine. Solche neuen Vakzine, die nach Verabreichung an den Patienten anti-Tumor-Immunität verleihen, sind beispielsweise bei N. Restifo, Current Opinion in Immunology 8, S. 658-663 (1996) oder Ying et al., Nature Medicine 5(7), S. 823 ff., (1999) beschrieben. Für Zwecke der Gentherapie können die vorstehenden Nukleinsäuresequenzen auch in Form von kolloidalen Dispersionen zu den Zielzellen transportiert werden. Dazu zählen beispielsweise Liposomen oder Lipoplexe (Mannino et al., Biotechniques 6 (1988), 682).

Weiter ist es bevorzugt, um eine Tumor-Immunität zu erzeugen, eine vorstehende Nukleinsäuresequenz in Antigen-präsentierende Zellen zu transfizieren und diese dem Patienten zu injizieren. Hierbei wird ein Plasmid in vitro in eine Antigen-präsentierende Zelle (APCs), z.B. Dendritische Zellen, gegeben, die dann Antigene produzieren und über HLA-Moleküle einzelne Peptide präsentieren. Dabei kann die Plasmid-DNA über verschiedene Wege in die Antigen-präsentierenden Zellen gebracht werden:

- (a) als nackte DNA, z.B. mittels "Gene Gun" oder Elektroporation;
- (b) als Lipid- oder Liposomen-verpackte DNA oder RNA (Nair et al., Nature Biotechnology 16, S. 364 ff. (1998));
- (c) mit einem Virus als Vektor (Kim et al., J. of Immunotherapy 20(4), S. 276-286 (1997));
- (d) mit einem Bakterium als Transportvehikel für den Expressionsvektor (Medina et al., Eur. J. Immunol. 29, S. 693-699 (1999); Guzman et al., Eur. J. Immunol. 28, S. 1807-1814 (1998))

Es kann auch das von einer vorstehenden Nukleinsäure kodierte Protein, das im Zusammenhang mit dem Vorhandensein einer malignen Tumorerkrankung steht, hergestellt werden. Das dazu bevorzugt angewendete Verfahren umfaßt die Kultivierung der vorstehend beschriebenen Wirtszellen unter Bedingungen, die die Expression des Proteins erlauben (vorzugsweise stabile Expression), und Gewinnung des Proteins aus der Kultur. Geeignete Verfahren zur rekombinanten Herstellung des Proteins sind allgemein bekannt (siehe beispielsweise Holmgren, Annu.Rev.Biochem. 54 (1985), 237; LaVallie et al., Bio/Technology 11 (1993), 187; Wong, Curr.Opin.Biotech. 6 (1995), 517; Romanos, Curr.Opin.Biotech. 6 (1995), 527; Williams et al., Curr. Opin. Biotech. (1995), 538; und Davies, Curr. Opin.Biotech. 6 (1995), 543. Auch geeignete Reinigungsverfahren (beispielsweise präparative Chromatographie, Affinitätschromatographie, beispielsweise Immunoaffinitätschromatographie, HPLC etc.) sind allgemein bekannt. In diesem Zusammenhang soll erwähnt werden, daß das vorstehend erwähnte Protein gemäß üblicher, auf dem Fachgebiet bekannten, Verfahren modifiziert sein kann. Zu diesen Modifikationen zählen

Austausche, Insertionen oder Deletionen von Aminosäuren, die die Struktur des Proteins modifizieren, wobei seine biologische Aktivität im wesentlichen erhalten bleibt. Zu den Austauschen zählen vorzugsweise "konservative" Austausche von Aminosäureresten, d.h. Austausche gegen biologisch ähnliche Reste, z.B. die Substitution eines hydrophoben Rests (z.B. Isoleucin, Valin, Leucin, Methionin) gegen einen anderen hydrophoben Rest, oder die Substitution eines polaren Rests gegen einen anderen polaren Rest (z.B. Arginin gegen Lysin, Glutaminsäure gegen Asparaginsäure etc.). Deletionen können zur Erzeugung von Molekülen führen, die eine deutlich geringere Größe aufweisen, d.h., denen beispielsweise Aminosäuren am N- oder C-Terminus fehlen.

Für die gewünschte Anti-Tumor-Vakzinierung eignen sich auch Injektionen mindestens eines der Proteins oder eines oder mehrerer davon abgeleiteter Peptide. Aus der Sequenz des erfindungsgemäßen Proteins werden dafür entweder mittels entsprechender Computerprogramme oder mittels Experimente (z.B. phagozytotische Aufnahme des Gesamt-Proteins, danach Analyse der präsentierenden Peptide) HLA-abhängige Peptid-Fragmente ermittelt. Diese werden mittels dem Fachmann bekannter Methoden künstlich hergestellt und dann (ggf. mit Immunsystem stimulierenden Faktoren, z.B. Interferonen, Interleukinen usw.) dem Patienten injiziert. Das hinter dieser Behandlung steckende Ziel ist, daß die APCs die Peptide aufnehmen, sie präsentieren und so in vivo die Produktion von Tumor-spezifischen zytotoxischen T-Zellen stimulieren. Allgemein ist dieses Prinzip von Melief et al., Current Opinion in Immunology 8, S. 651-657 (1996) beschrieben worden.

Ebenso wie vorstehend beschrieben, können anstelle des Vektors auch das vorstehende Protein bzw. Fragmente davon in vitro auf APCs geladen werden. Die beladenen Zellen werden dann dem Patienten z.B. in die Lymphknoten injiziert und sorgen direkt für die Stimulierung und Vermehrung von Tumor-spezifischen zytotoxischen T-Zellen (Nestle et al., Nature Medicine 4(3), S. 328 ff. (1998); Schadendorf et al., in: Burg, Dummer, Strategies for Immunointerventions in Dermatology, Springer Verlag, Berlin

Heidelberg, S. 399-409, 1997) Für die Vakzinierung kann es vorteilhaft sein, gegenüber dem Wildtyp-Antigen einzelne Aminosäuren, wie vorstehend beschrieben, zu modifizieren, da damit u.U. eine Erhöhung der Bindung und eine Verbesserung der Wirksamkeit zu erreichen ist (Clay et al., The Journal of Immunology 162, S. 1749-1755, 1999).

Besonders bevorzugt für die vorstehenden therapeutischen Maßnahmen ist mindestens eine Nukleinsäuresequenz, die die Nukleinsäuresequenz se20-10 (Fig.2), se57-1 (Fig.3), Lg1-2 (Fig. 5), se1-1 (Fig. 6), se2-1 (Fig.7), se2-2 (Fig.8), se14-3 (Fig.9), se20-7 (Fig.11), se20-9 (Fig.12), se33-1 (Fig.13), se37-2 (Fig.14), L14-2 (Fig.16), L15-7 (Fig.17), Li9-1 (Fig.18), Li9-4 (Fig.19), Lii5-2 (Fig. 20), Lii10-6 (Fig. 21), Liii4-5 (Fig. 22) oder GBP-TA umfaßt bzw. ein davon kodiertes Protein oder ein Fragment davon.

Die vorliegende Erfindung betrifft auch Antikörper, die die vorstehend beschriebenen Proteine (Tumor-Antigene) spezifisch erkennen. Die Antikörper können monoklonale, polyklonale oder synthetische Antikörper sein oder Fragmente davon, beispielsweise Fab-, Fv-oder scFv-Fragmente. Vorzugsweise handelt es sich dabei um monoklonale Antikörper. Für die Herstellung ist es günstig, Tiere, insbesondere Kaninchen oder Hühner für einen polyklonalen und Mäuse für einen monoklonalen Antikörper, mit einem vorstehenden (Fusions)protein oder Fragment(en) davon zu immunisieren. Weitere "Booster" der Tiere können mit dem gleichen (Fusions)protein oder Fragmenten davon erfolgen. Der polyklonale Antikörper kann dann aus dem Serum bzw. Eigelb der Tiere erhalten werden. Die erfindungsgemäßen Antikörper können gemäß Standardverfahren hergestellt werden, wobei das von den vorstehend erwähnten Nukleinsäuresequenzen kodierte Protein oder ein synthetisches Fragment davon als Immunogen dienen. Monoklonale Antikörper können beispielsweise durch das von Köhler und Milstein (Nature 256 (1975), 495) und Galfré (Meth. Enzymol. 73 (1981), 3) beschriebene Verfahren hergestellt werden, wobei Maus-Myelomzellen mit von immunisierten Säugern stammenden Milzzellen fusioniert werden. Diese Antikörper können beispielsweise zur Immunpräzipitation der vorstehend

diskutierten Proteine oder zur Isolierung verwandter Proteine aus cDNA-Expressionsbanken verwendet werden. Die Antikörper können beispielsweise in Immunoassays in Flüssigphase oder an einen festen Träger gebunden werden. Dabei können die Antikörper auf verschiedene Art und Weise markiert sein. Geeignete Marker und Markierungsverfahren sind dem Fachgebiet bekannt. Beispiele für Immunassays sind ELISA und RIA.

Weiter können die Antikörper neben ihrer diagnostischen Eignung auch therapeutisch eingesetzt werden. Dabei dient z.B. ein von den vorstehenden Nukleinsäuresequenzen kodiertes Protein als Target für bispezifische Antikörper. Hierzu wird auf Kastenbauer et al., Laryngorhinootologie 78(1), S. 31-35 (1999) und Cao et al., Bioconj. Chem. 9(6), S. 635-644 (1998) verwiesen. Die erfindungsgemäßen Antikörper eignen sich z. B. dafür, um in Tumoren überexprimiertes Antigen abzufangen und so das Tumorstadium zu hemmen, da es Hinweise darauf gibt, daß in manchen Fällen das Vorkommen von Tumor-Antigenen nicht nur das Vorhandensein von malignen Tumoren indikativ anzeigt, sondern aktiv das Tumorstadium fördert.

Desweiteren kann durch den Einsatz von Antisense-DNA (RNA) bzw. Ribozymen eine Inhibierung der Translation der vorstehenden Nukleinsäuresequenzen, deren Expression in Tumoren erhöht ist, erreicht werden und somit ein therapeutischer Effekt spezifisch auf diese Nukleinsäuresequenzen bzw. Gene ausgeübt werden. Es bilden sich in den entsprechenden Tumorzellen RNA/DNA-Hybride, die so die Transkription verhindern und - im Fall der Antisense-RNA - gleichzeitig einen Abbau der Hybride (und somit der RNA) durch RNase H bewirken (Scanlon et al., The FASEB Journal 9, S. 1288-1296, 1995)

Die vorliegende Erfindung betrifft somit ein Arzneimittel oder eine diagnostische Zusammensetzung, das (die) die vorstehend beschriebenen Nukleinsäuresequenzen, Vektoren, Proteine, Antikörper etc. oder Kombinationen davon enthält, bzw. deren Verwendung für die Diagnose und/oder Therapie. Vorzugsweise werden diese zur Diagnose oder Behandlung von malignen Tumorerkrankungen, insbesondere CTCL, verwendet. Bevorzugt ist

die Bereitstellung eines Vakzinierungsmittels, das wie vorstehend beschrieben, entweder auf der Nukleinsäuresequenz oder dem Protein/Peptid basiert. Die diagnostische Zusammensetzung eignet sich dabei einerseits um eine maligne Tumorerkrankung festzustellen, aber auch um eine Verlaufskontrolle durchzuführen, beispielsweise therapiebegleitend.

Die vorstehenden Arzneimittel enthalten gegebenenfalls zusätzlich einen pharmazeutisch verträglichen Träger. Geeignete Träger und die Formulierung derartiger Arzneimittel sind dem Fachmann bekannt. Zu geeigneten Trägern zählen beispielsweise Phosphat-gepufferte Kochsalzlösungen, Wasser, Emulsionen, beispielsweise Öl/Wasser-Emulsionen, Netzmittel, sterile Lösungen etc. Die Verabreichung der Arzneimittel kann oral oder parenteral erfolgen. Zu den Verfahren für die parenterale Verabreichung gehören die topische, intra-arterielle, intramuskuläre, subkutane, intramedulläre, intrathekale, intraventrikuläre, intravenöse, intraperitoneale oder intranasale Verabreichung. Die geeignete Dosierung wird von dem behandelnden Arzt bestimmt und hängt von verschiedenen Faktoren ab, beispielsweise von dem Alter, dem Geschlecht, dem Gewicht des Patienten, dem Stadium der Erkrankung, der Art der Verabreichung etc.

Eine vorstehende Nukleinsäuresequenz kann auch als Sonde verwendet werden, um DNA-Moleküle zu isolieren, die beispielsweise von einer anderen Spezies oder einem anderen Organismus stammen und ein Protein mit einer ebensolchen biologischen Aktivität kodieren. Vorzugsweise weist dazu die Sonde eine Länge von mindestens 20, besonders bevorzugt mindestens 25 Basen auf. Geeignete, auf Hybridisierung basierende Nachweisverfahren sind dem Fachmann bekannt, z.B. Southern oder Northern Blot. Geeignete Markierungen für die Sonde sind dem Fachmann ebenfalls bekannt und dazu zählen beispielsweise Markierung mit Radioisotopen, Biolumineszenz-, Chemilumineszenz-, Fluoreszenzmarkern, Metallchelaten, Enzymen etc..

Darüber hinaus kann dies auch durch eine PCR (Wiedmann et al., PCR Methods Appl. 3, S. 551-564 (1994); Saiki et al., Nature 324, S. 163-166 (1986)) oder "Ligase chain reaction" (LCR) (Taylor et al., Curr. Opin. Biotechnol. 6, S. 24-29 (1995); Rouwendal et al., Methods Mol. Biol. S. 149-156 (1996)) erfolgen, wobei die Primer von der Sequenz in den Figuren abgeleitet sind und wobei geeignete Primer (hinsichtlich Länge, Komplementarität zur Matritze, dem zu amplifizierenden Bereich etc.) vom Fachmann gemäß üblicher Verfahren entworfen werden können.

Außerdem betrifft die vorliegende Erfindung ein Verfahren zur Diagnose von malignen Tumorerkrankungen, in vitro, wobei die vorstehenden Nukleinsäuresequenzen oder Fragmente davon als Sonde verwendet werden.

Für das diagnostische Verfahren ist es bevorzugt die Nukleinsäuren und/oder Proteine in Form eines ELISA-Kits, Protein-Chips, Nukleinsäure-Chips oder einer mit DNA, RNA oder Protein beladener Membran bereitzustellen.

Bei oben genanntem Verfahren können dem Fachmann bekannte Methoden bezüglich der Präparation von DNA oder RNA aus biologischen Proben, des Restriktionsverdaus der DNA, der Auftrennung der Restriktionsfragmente auf nach Größe auftrennenden Gelen, beispielsweise Agarosegelen, der Herstellung und Markierung der Sonde und des Nachweises der Hybridisierung, beispielsweise über "Southern-Blot" oder in-situ Hybridisierung, angewandt werden.

Vorzugsweise handelt es sich bei diesem Diagnose-Verfahren um ein Verfahren, das die folgenden Schritte umfaßt:

- Isolierung von Nukleinsäure aus dem Patienten,
- Durchführung einer LCR oder PCR mit geeigneten Primern oder einer Hybridisierungsanalyse mit einer oder mehreren geeigneten Sonden basierend auf einer Nukleinsäuresequenz der Figuren,
- Nachweis eines amplifizierten Produkts oder einer Hybridisierung als Indiz auf das Vorliegen (oder

Nichtvorliegen) einer Tumorerkrankung (in Abhängigkeit davon, ob die jeweilige Nukleinsäuresequenz im Tumor im Vergleich zu Kontrollgewebe stärker oder schwächer (bzw. nicht) exprimiert wird.)

Dabei werden Primer verwendet, die eine vorstehend diskutierte Nukleinsäuresequenz oder geeignete Teilbereiche flankieren. Diagnostisch von Bedeutung sind dabei Amplifikationsprodukte von mRNA aus dem fraglichen Gewebe, die sich hinsichtlich des Auftretens von Tumor-spezifischen, insbesondere CTCL-spezifischen Banden von den Amplifikationsprodukten von mRNA aus gesundem Gewebe unterscheiden.

In einer alternativen bevorzugten Ausführungsform kann ein Verfahren angewendet werden, das folgende Schritte umfaßt:

- Isolierung von RNA aus dem Patienten,
- Durchführung einer Northern-Analyse mit einer oder mehreren geeigneten Sonden,
- Vergleich der Konzentration und/oder Länge der entsprechenden mRNA der Patientenprobe mit einer mRNA einer gesunden Person, wobei eine erhöhte bzw. erniedrigte Konzentration an mRNA (in Abhängigkeit von dem entsprechenden Marker; siehe Tabelle 4) im Vergleich zur Kontroll-mRNA aus Normalgewebe indikativ für eine Tumorerkrankung, insbesondere CTCL ist.

Bei diesem Verfahren können dem Fachmann bekannte Methoden bezüglich der Präparation von Gesamt-RNA bzw. poly(A)+RNA aus biologischen Proben, der Auftrennung der RNAs auf nach Größe auftrennenden Gelen, beispielsweise denaturierenden Agarosegelen, der Herstellung und Markierung der Sonde und des Nachweises über "Northern-Blot" angewandt werden.

In einer weiteren alternativen Ausführungsform kann eine mögliche Tumor-Erkrankung auch durch ein Verfahren diagnostiziert werden, das folgende Schritte umfaßt:

- Gewinnung einer Zellprobe von dem Patienten,
- Inkontaktbringen der so erhaltenen Zellprobe mit einem oder mehreren von den erfindungsgemäßen Nukleinsäuresequenzen

kodierten Proteinen oder Fragmenten davon als Sonde(n) unter Bedingungen, die die Bindung von Antikörpern erlauben, wobei die Gegenwart von Antikörpern in der Zellprobe indikativ für eine Tumorerkrankung, insbesondere CTCL ist.

Dieser Nachweis kann ebenfalls unter Anwendung von dem Fachmann bekannten Standardtechniken durchgeführt werden. Diesem sind auch Zellaufschlußverfahren bekannt, die die Isolierung der Antikörper auf eine solche Weise erlauben, daß diese mit dem Antigen in Kontakt gebracht werden kann. Der Nachweis des gebundenen Antikörpers erfolgt vorzugsweise über Immunassays, beispielsweise Western-Blot, ELISA, FACS oder RIA oder immunhistochemische Verfahren. Vorzugsweise erfolgt dieser über ELISA oder "Dot blot". Zur Etablierung eines ELISA können die vorstehenden Nukleinsäuresequenzen oder Fragmente davon in Expressionsplasmide kloniert und die entspr. Proteine rekombinant hergestellt werden, vorzugsweise als Fusionsproteine mit einem "His-Tag", was deren Aufreinigung erleichtert. Die Proteine werden dann auf Membranen oder andere geeignete Oberflächen aufgetragen, ggf. fixiert und mit adäquat verdünnten Patientenseren inkubiert. Nach den üblichen Waschschritten erfolgt dann eine Inkubation mit einem sekundären markierten Antikörper gemäß Routineverfahren zum Nachweis der gebundenen Patienten-Antikörper. Vorzugsweise werden die Patientenseren mit einer Vielzahl von Markerproteinen (Antigenen) inkubiert, da der Nachweis des Vorhandenseins (oder Fehlens) verschiedener Antikörper besser auf die zugrundeliegende Tumor-Erkrankung schließen läßt und evtl. auch eine Einteilung nach dem Erkrankungsstadium erlaubt.

Weiter betrifft die vorliegende Erfindung einen Kit zur Durchführung der erfindungsgemäßen Diagnoseverfahren, der den erfindungsgemäßen Antikörper oder ein Fragment davon, ein vorstehendes Protein (oder davon abgeleitetes Peptid), eine erfindungsgemäße Nukleinsäuresequenz (als Sonde) oder einen z.B. für PCR oder LCR geeigneten, auf der Sequenz der erfindungsgemäßen Nukleinsäuresequenzen basierenden Primer (oder ein Primerpaar) enthält, gegebenenfalls in Kombination mit einem geeigneten Nachweismittel.

Je nach Ausgestaltung des mit dem erfindungsgemäßen Kit durchzuführenden Diagnoseverfahrens können die in dem Kit enthaltenden Verbindungen (Nukleinsäuremoleküle, Proteine, Antikörper oder Fragmente davon) an einem geeigneten Träger immobilisiert sein, d.h. in Form eines Chips oder auf einer Membran gebunden vorliegen.

Alle vorstehend erwähnten Proteine werden serologisch nur von Antikörpern aus Seren von Tumorpatienten, nicht aber von Seren aus Kontrollpersonen, erkannt und sind somit serologisch tumorspezifisch. Da sich diese Spezifität nicht auf einen Tumortyp beschränkt, eignen sich diese Proteine und Antikörper sehr gut, um überhaupt eine Unterscheidung zwischen Malignität und "Nicht-Malignität" zu treffen. Es hat sich dabei als vorteilhaft herausgestellt, die Untersuchung mit mehr als einem der vorstehend erwähnten Tumormarker, d.h. einer Kombination von Tumormarkern, durchzuführen und abhängig davon einen therapeutischen Ansatz zu wählen. Dies bedeutet, daß auch das erfindungsgemäße Arzneimittel mehr als eine der vorstehend erwähnten Nukleinsäuren, Proteine oder Antikörper enthalten sollte.

Die Erfindung wird nun weiter anhand der Figuren beschrieben, welche zeigen:

Fig.1: Nukleinsäuresequenz von se2-5 und ein davon abgeleiteter ORF

Fig.2: Nukleinsäuresequenz von se20-10 und ein davon abgeleiteter ORF

Fig.3: Nukleinsäuresequenz von se57-1 und ein davon abgeleiteter ORF

Fig.4: Nukleinsäuresequenz von se70-2 und ein davon abgeleiteter ORF

Fig. 5: Nukleinsäuresequenz von Lg1-2 und ein davon abgeleiteter ORF

- Fig.6: Nukleinsäuresequenz von se1-1 und ein davon abgeleiteter ORF
- Fig.7: Nukleinsäuresequenz von se2-1 und ein davon abgeleiteter ORF
- Fig.8: Nukleinsäuresequenz von se2-2 und ein davon abgeleiteter ORF
- Fig.9: Nukleinsäuresequenz von se14-3 und ein davon abgeleiteter ORF
- Fig.10: Nukleinsäuresequenz von se20-4 und ein davon abgeleiteter ORF
- Fig.11: Nukleinsäuresequenz von se20-7 und ein davon abgeleiteter ORF
- Fig.12: Nukleinsäuresequenz von se20-9 und ein davon abgeleiteter ORF
- Fig.13: Nukleinsäuresequenz von se33-1 und ein davon abgeleiteter ORF
- Fig.14: Nukleinsäuresequenz von se37-2 und ein davon abgeleiteter ORF
- Fig.15: Nukleinsäuresequenz von se89-1 und ein davon abgeleiteter ORF
- Fig.16: Nukleinsäuresequenz von L14-2 und ein davon abgeleiteter ORF
- Fig.17: Nukleinsäuresequenz von L15-7 und ein davon abgeleiteter ORF
- Fig.18: Nukleinsäuresequenz von Li9-1 und ein davon abgeleiteter ORF

Fig.19: Nukleinsäuresequenz von Li9-4 und ein davon abgeleiteter ORF

Fig.20: Nukleinsäuresequenz von Lii5-2 und ein davon abgeleiteter ORF

Fig.21: Nukleinsäuresequenz von Lii10-6 und ein davon abgeleiteter ORF

Fig.22: Nukleinsäuresequenz von Liii4-5 und ein davon abgeleiteter ORF

Fig. 23: Nukleinsäuresequenz von GBP-TA und ein davon abgeleiteter ORF

Fig. 24: Lokalisation von GBP-TA, GBP-TA_{short} und Lg1-2 auf dem Chromosom 1p22.3. Die verwendeten Primer zur Unterscheidung der Splicing-Varianten sind eingezeichnet.

Primerset I: tgt tgt aga tca ctt caa ggt gc (forw.)
cca tat cca aat tcc ctt ggt gtg ag (re.)
Annealing-Temp. 63°C

Primerset II: aga agg aag aaa ctc caa aca cat cc
(forw.)
cca tat cca aat tcc ctt ggt gtg ag
(re.)
Annealing Temp. 48°C

Die Erfindung wird nun nachfolgend mit Bezug auf die Beispiele beschrieben.

Hinsichtlich der verwendeten Methoden wird neben den in Beispiel 1 angegebenen Verfahren außerdem auf Sambrook, J, Fritsch, E.F. und Maniatis, T. (Molecular cloning; a laboratory manual; second edition; Cold Spring Harbor Laboratory Press, 1989) und Current Protocols in Molecular Biology (John Wiley and Sons, 1994-1998) hingewiesen, wobei die nachfolgend erwähnten Techniken,

insbesondere Screenen von cDNA-Bibliotheken, Präparation von DNA bzw. RNA, PCR, RT-PCR oder Northern Blot dem Fachmann hinreichend bekannt sind und beherrscht werden.

Beispiel 1: Allgemeine Verfahren

(A) Gewebe und Seren

Seren und Tumorgewebe wurden bei diagnostischen oder therapeutischen Routineverfahren mit Einverständnis der Patienten (und der Erlaubnis des zuständigen Ethikkomitees) erhalten. Die Lagerung der Gewebe und Seren erfolgte bei -20°C bzw. -80°C.

(B) Herstellung der von cDNA-Banken

mRNA wurde aus Testis-Proben unter Verwendung eines Kits für RNA-Isolierung ("RNeasy midi kit"; Qiagen, Hilden, Deutschland) und danach eines Kits für mRNA-Isolierung ("oligotex mRNA kit"; Qiagen) entsprechend den Empfehlungen des Herstellers extrahiert. Für die Konstruktion der λ -ZAP-Expressionsbank ("UNI-ZAP[™] XR custom cDNA library"; Stratagene, La Jolla, CA, USA) wurden insgesamt 10,4 μ g mRNA verwendet. Die cDNA-Bank bestand aus 10^6 primären Rekombinanten mit einer Insertionslänge über 0,4 kbp und wurde zu 10^{10} Plaque-bildenden Einheiten (pfu) amplifiziert. Für die Konstruktion der CTCL-Bank wurden 4,8 μ g mRNA aus verschiedenen Proben von kutanen T- und B-Zell Lymphomen verwendet. Die Zahl der primären Rekombinanten betrug 6×10^7 . Das Vorgehen war analog zum vorbeschriebenen Vorgehen bei der Testis-Bank.

(C) Immunscreening

Das Immunscreening wurde wie in Sahin et al. (PNAS USA 92 (1995), 11810-11813) und Türeci et al. (Cancer Res. 56 (1996), 4766-4772) beschrieben durchgeführt. Alle Seren wurden in Tris-gepufferter Saline (TBS mit 0,2% Trockenmilchpulver, pH-Wert

7,5) verdünnt und mit E.coli-Proteinen (mechanisch aufgebrochen oder durch Phagen ohne Insertion lysiert) vorabsorbiert. Mit rekombinanten λ -ZAP-Phagen transduzierte E.coli wurden auf NZY-Agar bei einer Konzentration von 2000 Plaques/Platte ausplattiert und die Expression der rekombinanten Proteine wurde mittels Isopropyl- β -D-Thiogalactosid induziert. Die Platten wurden bei 37°C ü.N. inkubiert und die Proteine 4 Stunden bei 37°C auf Nitrozellulose-Membranen transferiert und gebunden. Die Membrane wurden mit Tween-20™ (0,05%) enthaltender TBS gewaschen, mit 5% Trockenmilchpulver in TBS abgesättigt und mit Seren (entweder von Patienten oder als Kontrolle von gesunden Personen) bei einer Endkonzentration von 1/100 inkubiert. Reaktive Proteine wurden mit einem Alkalische-Phosphatase-gekoppelten sekundären Antikörper (Ziege-anti-human-IgC, Fc-Fragment; Dianova, Hamburg, Deutschland) nachgewiesen und mittels 5-Brom-4-Chlor-3-Indolylphosphat und Nitroblau-Tetrazolium sichtbar gemacht. Positive Phagemids wurden mittels Seren von Patienten mit Mycosis fungoides (n=15) und Sézary Syndrom (n=3) sowie gesunden Personen als Kontrolle (n=10) weiter untersucht. Positive Phagemids wurden zur Monoklonalität subkloniert und einer in vivo-Exzision des "pBluescript"-Plasmids (entsprechend dem Protokoll des Herstellers der Genbank, Stratagene, La Jolla, CA, USA) unterzogen. DNA wurde entsprechend dem Protokoll des Herstellers isoliert ("QIAprep spin miniprep"; Qiagen). Die Größe der Insertionen wurde durch eine SmaI/KpnI-Spaltung und Gelelektrophorese analysiert. Die Sequenzierung wurde mittels eines automatischen Fluoreszenz-Sequenziergeräts (Modell 377; Perkin-Elmer / Applied Biosystems, Forster System, CA, USA) und des Dye-Terminator-Verfahrens entsprechend den Angaben des Herstellers ("ABI PRISM Big Dye Ready Reaction Terminator Cycle Sequencing Kit"; Perkin-Elmer) durchgeführt. Primer wurden chemisch synthetisiert. Die Sequenzen der Klone wurden auf beiden komplementären Strängen vollständig ermittelt.

(D) Tumorgewebe und Zelllinien

Von 17 CTCL-Patienten erhaltene Gewebeproben dienten als Quelle für die herstellung von Tumor-cDNA: 13 Mycosis fungoides

(Stadium Ib bis IVb, hauptsächlich IIb), 2 Sézary Syndrome (Stadium III), 1 T-Zonen-Lymphom (Stadium IVb) und 1 CD30+ CTCL (Stadium IIb). Außerdem wurden cDNAs von den folgenden 4 CTCL-Zelllinien hergestellt: My-La (Mycosis fungoides; Kaltoft et al., In Vitro Cell Dev. Biol. 28a (1992), 161-167), SeAx (Sézary Syndrom; Kaltoft et al., Arch. Dermatol. Res. 280 (1988), 264-267), HH (Lymphomatoide Papulose; ATCC-Nr.: CRL-2105) und HuT-78 (Sézary Syndrom; ATCC-Nr.: TIB-161). Außerdem wurde cDNA von sechs Leukämie-Zelllinien (ARA-10, Jurkat, KG1, K562, Nalm-2 und SKW6.4 und 22 Melanom-Zelllinien hergestellt.

Zur Analyse der Gewebeverteilung innerhalb normalen Geweben wurden ausführlich Kontroll-cDNAs verwendet, dazu zählten auch drei Felder von im Handel erhältlichen cDNAs (alle von Clontech, CA, USA): humane "multiple tissue"-cDNA Feld I, Feld II und humanes fötales MTC-Feld. Zusätzlich wurden verschiedene im Handel erhältliche Gesamt-RNAs zur Herstellung weiterer Kontroll-cDNAs mittels des vorstehend beschriebenen Verfahrens erzeugt. Schließlich dienten auch noch cDNAs von drei aktivierten CD8+ T-Zelllinien (Möller et al., British Journal of Cancer 77 (1998), 1907-1916) als Kontroll-T-Zellen.

(E) RT-PCR

Aufgrund der eingeschränkten Menge an RNA wurde bevorzugt RT-PCR zur Untersuchung der identifizierten Sequenzen innerhalb verschiedener normaler Gewebe und Tumor-Gewebe verwendet. In ausgewählten Fällen wurden diese Untersuchungen mittels Northern-Blot-Analysen vervollständigt. RT-PCR wurde mittels "MJ Research PCT-200" (Biozym, Oldendorf, Deutschland) mit einer einminütigen Anlagerung bei variabler Temperatur mit 35 Zyklen durchgeführt. Alle RT-PCR wurden in mindestens zwei unabhängigen Experimenten durchgeführt. Die RNA-Isolierung, RT-PCR und die Northern-Blot-Analysen wurden ansonsten gemäß üblicher Standardverfahren und Standardbedingungen durchgeführt. Die Primersequenzen und Annealing-Temperaturen für die verschiedenen Klone sind in der nachfolgenden Tabelle 1 angegeben:

Klon	Forward Primer	Reverse Primer	Anneal. Temp.
se1-1	gca aaa gca att aga cgc tac c	cac agc cct gtt ctt ctt tag c	57°C
se2-1	gta cag cag aaa gca agc aac tga atg	gga aat tgg att cta aag cag ttc ctt c	55°C
se2-2	cta tga atc caa gac caa agg c	ctc cac ttt ggt cct tgt tag c	59°C
se2-5	acc cac gca gat ttg gaa tc	agg ctg atc act ggc tgt g	59°C
se14-2	cct tat tgt aca ctg ggg ctt c	cag aca caa gga act gaa gta acg	60°C
se14-3	cac tgc caa gat aga caa gca g	gct ctt atc cag gaa gtc cat g	59°C
se20-4	tac agg atc tca gac ata tct cca tg	aaa tgt ctt ccc act gca taa tag tc	59°C
se20-7	taa gga aac aat tca gtc aca taa gg	ctg tag ctt agc aat ttg ttct tot g	59°C
se20-9	tta tga ggc tta gaa ttt caa cca c	aaa ggc ttt caa aac att ttt caa c	59°C
se20-10	gta gag atc aga gag ttg tga cat ctg	tat tac ttt tca ctg tta cac tgc tgg	59°C
se33-1	gcc aca gag aat gaa cca ctt aac	gag gga cta tca gti got gti tg	60°C
se37-2	gca tct aat aga acg cta cta cca cc	ctg tga gct atc acc tat cct tga g	60°C
se57-1	gtg aca gtg acc aca gaa att ccc cc	cac gtt tct cag agc tgc tgc tcc	63°C
se70-2	gct gca cag aaa acc tta ctt gti tcc acc	ctc gta aat gca gaa atc tcc aat gcc c	56°C
se89-1	tcc aca gcc tat tgg ctc act tgg ac	gcc ctt tag tgt gtc tgt aat tgg aat cag	57°C
RAP140	tcc aca gcc tat tgg ctc act tgg ac	gca cac act gct cct cca cct gac	57°C
L14-2	got got gct gti tac aga aag got cac	gga aag tta tcc aca gct act gag gao cc	64°C
L15-7	tcc cct cca ttt aat ctc caa att cac cc	ctc agc att tgc cgc cgt aac tt	62°C
Li9-1	gaa aac tac aaa tcc cag gag cac	ctc acg aaa tat gag ctt cac cac	63°C
Li9-4	tta ctg atc gtc tgc tcc cta gag tcc	atc ttc tgc tca gtc aga atc cca tgo	67°C
Lgl-2 = Primer set I (GBP-TA)	tgt tgt aga tca ctt caa ggt gc	cca tat cca aat tcc ctt ggt gtg ag	63°C
GBP-TA Primer set II	tgt tgt aga tca ctt caa ggt gc	cca tat cca aat tcc ctt ggt gtg ag	48°C
Lii5-2	tga gaa tga ggt ggg ggt gg	tgg gga acc gga tca gga c	58°C
Liii10-6	gca tcc tac cac caa ctc gtc c	agt tct gag acc gti ctt cca cc	57°C
Liiii4-5	gct gcg gac ata aat ctt aaa gc	agg gtc tca ctc tga ttg cc	56°C

(F) Northern Blot

10 µg Gesamt-RNA werden auf einem MOPS-Gel elektrophoretisch aufgetrennt und auf positiv geladene Nylon-Membranen übertragen. Die Sondenmarkierung erfolgt mittels des Roche High Prime Kits mit α -³²P-dCTP. Die Entfernung der nicht eingebauten Nukleotide erfolgt mittels des Qiagen Removal Kits (Fa. Qiagen, Hilden). Die Prähybridisierung wird bei 60°C für 1-2 Stunden durchgeführt. Die Hybridisierung erfolgt bei 60°C vorzugsweise über Nacht. Die Prähybridisierungs- und die Hybridisierungs-Lösungen haben die Zusammensetzung: 10% Dextransulfat, 1% SDS, 10x Denhardts Reagenz, 3x SSC. Das nachfolgende Waschen erfolgt für 2 x 30 Minuten mit 2xSSC/0,1% SDS bei 42°C und dann 2 x 30 Minuten mit 0,2xSSC/0,1% SDS bei 65°C. Es wird ein Kodak X-Omat-Film für eine Expositionszeit von 3-10 Tagen aufgelegt.

Beispiel 2: Screenen nach positiven Klonen

Ungefähr $1,9 \times 10^6$ rekombinante Klone einer von normalem Testis-Gewebe erhaltenen cDNA-Bank wurden mit Seren von Patienten mit kutanen T-Zell-Lymphomen (CTCL) einschließlich Mycosis fungoides (MF) und Sézary Syndrom gescreent. Es konnten 28 Klone, die 22 verschiedene ORFs bzw. Gene repräsentieren, nachgewiesen werden und diese wurden hinsichtlich serologischer Reaktivität und molekularer Verteilung weiter untersucht. Eine sekundäre Bestätigung wurde durch die Verwendung zusätzlicher Seren von Patienten mit einer positiven Diagnose für Mycosis fungoides (MF) (n=15) oder Sézary Syndrom (n=3) und 10 Kontrollseren von gesunden Freiwilligen durchgeführt. Die Reaktivität der Seren der Patienten war mit dem Tumorstadium (höchstens Stadium III) assoziiert (Tabelle 2). Dies war jedoch nicht statistisch signifikant (χ^2 -Test und Mann-Whitney U-Test), vermutlich aufgrund der niedrigen Anzahl von Seren mit hohen Tumorstadien. Die Reaktivität der Patienten-Seren war im Bereich von 11% bis 71% von rekombinante Klone identifizierenden Seren.

Tabelle 2: Anzahl positiver Klone in Korrelation mit dem Tumorstadium des Serum-Spenders

Tumor-Stadium des Patienten	Primär-Screen			Sekundär-Screen		
	Anzahl der eingesetzten Seren	gescreente Plaques (positiv / gesamt) ⁽¹⁾	Häufigkeit ⁽²⁾	Anzahl der eingesetzten Seren	Plaques (positiv / gesamt)	Häufigkeit ⁽²⁾
I	6	1 / 524.000	$0,2 \times 10^{-05}$	7	10 / 106	0,09
II	6	15 / 790.000	$1,9 \times 10^{-05}$	6	28 / 84	0,33
III	3	12 / 274.000	$4,4 \times 10^{-05}$	3	19 / 47	0,40
IV	2	0 / 270.000	0	2	11 / 39	0,28
gesamt ¹	17	28 / 1.858.000	$1,5 \times 10^{-05}$	18	68 / 276	0,25

Primär Screen einer Testis cDNA Bank wurde nacheinander mit 17 einzelnen Seren durchgeführt, die von Patienten stammten, die Tumoren im angegebenen Stadium aufwiesen. Die Anzahl der positiven und gesamten analysierten Plaques sind in der 3. Spalte angegeben. Während des *Sekundär Screens* wurde jeder positive Plaque (28) mit bis zu 18 individuellen Seren verschiedener Patienten im angegebenen Tumorstadium nachgetestet.

⁽¹⁾ Kumuliert über alle getesteten Seren. ⁽²⁾ Positive Klone dividiert durch Gesamtzahl der getesteten Klone.

Die Anzahl und Wahrscheinlichkeit aller serologischen Antworten sind hinsichtlich des Tumorstadiums der Patienten zusammengefaßt, von denen Serum für das Screenen einer Testis-cDNA-Bank genommen wurde. Obwohl die Daten eine höhere Wahrscheinlichkeit für das Vorhandensein von Antikörpern gegen Tumor-Antigene (mit dem Gipfel bei Stadium III) anzeigen, sind diese Unterschiede nicht statistisch signifikant.

Tabelle 3: Serologische Analysen der identifizierten Klone

Klon	CTCL		Kontrollen	
	reaktiv	n	reaktiv	n
se1-1	50%	10	0%	5
se2-1	22%	18	0%	10
se2-2	33%	9	0%	8
se2-5	30%	10	0%	5
se14-3	11%	9	0%	5
se20-4	40%	10	0%	5
se20-7	30%	10	0%	7
se20-9	25%	8	0%	5
se20-10	11%	18	0%	10
se33-1	29%	17	0%	10
se37-2	29%	17	0%	10
se57-1	33%	15	0%	9
se70-2	10%	10	0%	6
se89-1	44%	18	0%	9
L14-2	42%	19	0%	5
L15-7	33%	21	0%	7
Li9-1	19%	21	0%	5
Li9-4	57%	21	0%	6
Lg1-2	56%	16	0%	9
Lii5-2	20%	15	0%	8
Lii10-6	6%	16	0%	8
Liii4-5	29%	17	0%	8

Die Tabelle gibt die prozentuale Reaktivität der Seren (Anzahl n) gegen die getesteten Klone im sekundären Screen an.

Der Prozentsatz der reagierenden und die Gesamtzahl der getesteten Seren (n) während des sekundären Screenens sind angegeben. Bei den Klonen se2-1 und se20-4 ist zusätzlich jeweils einer der homologen Klone angegeben, da diese sich in ihrem Reaktionsmuster unterschieden, vermutlich aufgrund von Sequenzunterschieden.

Fünf Antigene repräsentieren bisher noch nicht bekannte Sequenzen (se2-5, se20-10, se57-1, se70-2 und Lg1-2). Das mittels RT-PCR analysierte RNA-Expressionsmuster der identifizierten Antigene variierte zwischen hocheingeschränkter und ubiquitärer Expression in 28 normalen, 17 CTCL-Tumorgeweben und 33 Tumorzelllinien unterschiedlichen Ursprungs (Tabellen 4 und 5).

Tabelle 4: Expressionsanalyse mittels RT-PCR mit Antigen-spezifischen Primern und cDNA von verschiedenen Geweben und Zell-Linien

<i>cDNA</i>	<i>Kontrollen</i>		<i>Tumor Gewebe</i>	<i>Zell-Linien</i>		
<i>Primer gegen Klon</i>	<i>multi-tissue panels ⁽¹⁾ (n)</i>	<i>aktivierte CTLs ⁽²⁾ (n=3)</i>	<i>CTCL (n)</i>	<i>CTCL (n=4)</i>	<i>Leukämie (n=5-6)</i>	<i>Melanom (n)</i>
se1-1	100% (17)	100%	91% (11)	100%	100%	100% (6)
se2-1	4% (28) ^(1,3)	0%	6% (17)	0%	0%	0% (11)
se2-2	100% (18)	100%	90% (10)	100%	83%	80% (5)
se2-5	94% (18)	0%	55% (11)	0%	100%	100% (11)
se14-3	100% (18)	100%	100% (11)	100%	100%	100% (11)
se20-4	100% (18)	100%	92% (12)	100%	100%	100% (11)
se20-7	100% (15)	100%	100% (11)	100%	50%	80% (5)
se20-9	100% (18)	100%	82% (11)	100%	83%	45% (11)
se20-10	46% (28) ⁽³⁾	67%	77% (13)	100%	100%	55% (11)
se33-1	61% (28) ⁽³⁾	100%	75% (16)	100%	83%	100% (11)
se37-2	100% (16)	100%	93% (15)	100%	83%	75% (5)
se57-1	21% (28) ⁽³⁾	0%	6% (17)	0%	0%	0% (23)
se70-2	54% (28) ⁽³⁾	33%	31% (16)	100%	100%	45% (22)
se89-1	87% (23)	100%	75% (16)	100%	100%	73% (22)
L14-2	75% (24)	33%	40% (15)	50%	60%	10% (21)
L15-7	65% (26)	100%	33% (15)	50%	40%	20% (20)

Li9-1	85% (26)	33%	80% (15)	75%	100%	42% (12)
Li9-4	92% (26)	100%	73% (15)	75%	80%	60% (15)
GBP-TA	32% (28) ^(3,4)	0%	26% (19)	75%	20%	0% (23)
Lii5-2	59% (22)	0%	21% (14)	25%	75%	21% (24)
Lii10-6	86% (21)	100%	100% (14)	50%	100%	nt
Liii4-5	55% (22)	33%	29% (14)	100%	75%	nt

Die Tabelle gibt die prozentuale Häufigkeit der Expression in verschiedenen Geweben und Zell-Linien nach RT-PCR-Analyse an (Anzahl der Gewebe ist in Klammern; nt: nicht getestet). ⁽¹⁾ RT-PCRs an Testis cDNA ergab immer positive Ergebnisse. Im Falle von Klonse2-1 war Testis cDNA die einzige positive Probe. Zusammensetzung der Kontroll-Palette siehe Table 5. ⁽²⁾ Aktivierte zytotoxische T-Zellen. ⁽³⁾ Details zu diesen Ergebnissen siehe Table 5. ⁽⁴⁾ GBP-TAshort mit 7%.

Tabelle 5: RT-PCR-Analysen mittels spezifischen Primern gegen differentiell exprimierte Sequenzen und "multiple tissue" (MTC) - cDNA

Primer gegen Klon:	se2-1	se20-10	se33-1	se57-1	se70-2	GBP-TA/ -TAshort
Darm	-	-	-	+	-	-
Dünndarm	-	+	+	+	+	-/+
fetale Leber	-	-	-	-	+	-/+
fetale Lunge	-	-	+	-	+	-
fetale Milz	-	-	+	-	+	-/+
fetale Niere	-	+	+	-	+	-
fetaler Skelettmuskel	-	-	+	-	+	-
fetaler Thymus	-	-	-	-	+	-/+
fetales Gehirn	-	-	+	-	+	-
fetales Herz	-	-	+	-	+	-/+
Gehirn	-	+	+	-	+	-
Haut	-	-	+	-	-	-
Herz	-	+	-	-	-	-
Knochenmark	-	+	+	+	-	+
Leber	-	-	+	-	-	-
Lunge	-	-	+	-	-	-

WO 02/38803					PCT/DE01/04229	
Magen	-	+	+	-	+	
Milz	-	+	+	+	+	-/+
Niere	-	+	+	-	+	-
Ovar	-	-	+	-	-	-
Pancreas	-	-	+	-	-	-
periph. Blut Lymphozyten	-	-	-	-	-	-
Placenta	-	+	+	-	-	-/+
Prostata	-	+	+	-	-	-
Skelettmuskel	-	-	-	-	-	-
Testis	+	+	+	+	+	-
Thymus	-	-	+	-	+	-
Trachea	-	+	+	+	+	-

Es wurden entweder kommerziell erhältliche cDNA Paletten verwendet oder cDNA aus RNA paletten hergestellt (Clontech). Jede Probe enthielt cDNA aus Proben mehrerer Individuen. Die Haut cDNA wurde aus einer einzelnen Probe hergestellt. RT-PCR gegen GBP-TA / GBP-TAshort unterscheidet die beiden Varianten.

Beispiel 3: Tumorspezifische Antigene

Sechs Klone (die durch mindestens vier verschiedene Rekombinanten repräsentiert werden) waren zu SCP-1 homolog, einem mit der Meiose in Zusammenhang stehenden Protein (Türeci et al., PNAS USA 95 (1998), 5211-5216). Interessanterweise unterschied sich die serologische Reaktivität verschiedener Seren zwischen den unterschiedlichen SCP-1-Klonen: Klon se2-1 wurde durch Seren von 2/15 MF-Patienten und 2/3 von Patienten mit Sézary Syndrom nachgewiesen. Mittels RT-PCR wurde nachgewiesen, daß se2-1 tumorspezifisch ist. Ein weiterer zu SCP-1 homologer Klon (se37-1) wurde durch 3/9 MF-Seren und 1/3 Sézary Syndrom-Seren nachgewiesen. Dies könnte verschiedene Epitope von SCP-1 widerspiegeln, da die Klone sich in ihrer Länge unterschieden. Klon se33-2 reagierte außerdem auch mit 1/5 Kontrollseren. Interessanterweise kodierte dieser Klon eine weitere Peptidsequenz innerhalb des ersten Leserahmens, die innerhalb der anderen zu SCP-1 homologen Klone nicht vorhanden war und somit ein Autoantigen darstellen könnte, das für die Reaktivität des Kontrollserums verantwortlich ist. Die PCR-Analysen wurden mit den gleichen Primern wie von Türeci et al. (1998) veröffentlicht durchgeführt, die auch zu den SCP-1 homologen Klonen perfekt paßten. Innerhalb aller getesteten

normalen Geweben sowie den Tumorproben und Zelllinien konnten nur eine Testis-Probe und eine MF-Probe (Patient H.S.) gefunden werden, die SCP-1 mRNA exprimierte. Das positive Ergebnis der MF-cDNA konnte durch Northern-Blot bestätigt werden, der eine Bande von etwa 4,3 kb ergab. Außerdem reagierte das Serum des Patienten H.S. auch mit se2-1 und einem der anderen zu SPC-1 homologen Klone. Desweiteren sind gemäß Northern Blot Analyse die Klone se57-1 und L15-7 Tumor-spezifisch.

Beispiel 4: (A) Antigene mit eingeschränktem Expressionsmuster und (B) ubiquitär exprimierte Antigene

Nachfolgend werden 13 Antigene mit differentieller oder ubiquitärer Expression (nachgewiesen über RT-PCR) beschrieben (siehe Tabellen 4 und 5).

(A)

Für fünf neue Antigene (se2-5, se20-10, se57-1, se70-2 und Lg1-2) und zwei Antigene mit Homologien zu bekannten Genen (se33-1: NP220; se89-1: mit Retinoblastom in Zusammenhang stehendes Protein RAP140) zeigte sich eine differentielle Expression auf molekularer Ebene. Die serologische Reaktivität gegen diese Klone (definiert als Prozentsatz reaktiver Seren während des sekundären Screenens) betrug im Durchschnitt 31%. Eine geringe Reaktivitätsrate zeigte sich gegenüber den Klonen se20-10 und se70-2 (2/18 bzw. 1/18 reaktive Seren), während 71% der Seren von CTCL-Patienten (n=14) mit dem Klon se89-1 positiv reagierten. Alle 6 Klone zeigten mit bis zu 10 Kontrollseren keine Reaktion.

Es wurde begonnen, die RT-PCR-Ergebnisse mittels Northern Analysen zu quantifizieren. Solche Antigene, die sich in Normal-Geweben im Vergleich zu Tumorproben nicht als gleich stark exprimiert erweisen, werden als potentielle Therapeutika eingestuft.

Mittels RT-PCR-Analysen konnte gezeigt werden, daß die se-2-5-spezifische mRNA beinahe ubiquitär innerhalb normaler Gewebe exprimiert wird, jedoch nicht in aktivierten T-Zellen und nur in lediglich 55% der CTCL-Gewebeproben. Im Gegensatz dazu ergab die Northern-Blot-Analyse selbst innerhalb normaler Gewebe ein eingeschränktes Expressionsmuster. Starke Signale waren in Niere, Luftröhre und Testis nachweisbar, schwächere in Kolon, Dünndarm, Thymus, Knochenmark und Magen. Während in allen positiven normalen Geweben drei Banden nachweisbar waren (5,2, 4,2 und 3,9 kb) zeigte die einzige positive CTCL-Zelllinie SeAx ein Signal bei 3,9 kb.

Spezifische mRNAs für die Klone se20-10 und se57-1 wurden mittels RT-PCR in 43% bzw. 21% der untersuchten Kontrollgewebe gefunden. Interessanterweise war die Expression der für se57-1 spezifischen mRNA in allen Tumorgeweben und Zelllinien sehr stark herunterreguliert. Im Gegensatz dazu war die Expression der mRNA von Klon se70-2 im Vergleich zu normalen Geweben (54%) innerhalb der CTCL- und Leukämie-Zelllinien hochreguliert (100%), während die CTCL-Gewebe und Melanom-Zelllinien mittlere Expressionsspiegel zeigten (33% bzw. 45%). Unter den Kontrollgeweben waren alle fötalen Gewebe in der RT-PCR positiv.

Für zwei Antigene mit Homologien zu bekannten Sequenzen zeigte sich eine differentielle Expression: se33-1 cDNA ist homolog zu NP220, einem DNA-bindenden Protein, und se89-1 cDNA ist homolog zu RAP140, einem Retinoblastom-assoziierten Klon. Klon se33-1 zeigte innerhalb einer überlappenden Strecke von 3830 bp am 3'-Ende von NP220 99% Ähnlichkeit, dieser Klon ist jedoch am 5'-Ende verkürzt, was zu einem verkürzten ORF führt. Die RT-PCR (unter Verwendung se33-1-spezifischer Primer) erbrachte den Nachweis von mRNA in 6/8 fötalen und 16/20 normalen Geweben (Tabelle 5) sowie innerhalb von CTCL-Geweben (12/16), aktivierten zytotoxischen T-Zellen und in den meisten Zelllinien (Tabelle 4).

cDNA des Klons se89-1 zeigte 98% Ähnlichkeit zu RAP140 innerhalb

eines überlappenden Bereichs von 3444 bp und eine Lücke von 60 bp, die innerhalb des ORF liegt und zu einer Aminosäure-Lücke von 20 Aminosäuren führt. RT-PCR wurde mit verschiedenen Primern durchgeführt: Zuerst mit der Primerkombination RAP140 (siehe Tabelle 1 sowie Beispiel 1), wobei sowohl RAP140 als auch se89-1 nachgewiesen wurden und sich drei Banden in Testis-cDNA und zwei Banden in verschiedenen anderen cDNAs zeigten. Zur Spezifizierung der Expression von se89-1 wurde ein neuer reverser Primer (s. Tabelle 1) entworfen, der die Lücke innerhalb se89-1 überspannt. Unter Verwendung dieses Primers zusammen mit dem Vorwärts-Primer gegen RAP140, der ebenfalls se89-1 nachweist, wurde nur eine Bande amplifiziert. Die Häufigkeit der positiven cDNAs zu den se89-1-spezifischen Primer unterschied sich nicht wesentlich zwischen den Kontrollgeweben (79%), CTCL-Geweben (75%) und Zelllinien (CTCL- und Leukämie-Zelllinien: 100%, Melanom-Linien: 73%). Mittels Northern-Blot-Analysen konnte die Gegenwart von für se89-1 spezifischer mRNA innerhalb von mRNA bestätigt werden, die von Hirn, Niere, Kolon und Testis und der CTCL-Linie SeAx stammte.

(B)

Sechs zu bekannten Sequenzen homologe Antigene (se1-1, se2-2, se14-3, se20-4, se20-9 und se37-2) erwiesen sich als ubiquitär exprimiert. In allen Kontrollgeweben (n=28) war für diese Klone spezifische mRNA über RT-PCR nachweisbar. In den zwei Klonen se14-3 und se20-4 waren alle Tumorgewebe und Zelllinien ebenfalls in der RT-PCR positiv. Im Gegensatz dazu wurden die Klone se2-2, se20-7, se20-9 und se37-2 nur in einer Untergruppe der Melanom- und Leukämie-Zelllinien exprimiert, während CTCL-Gewebe und CTCL-Zelllinien einen höheren Prozentsatz von in der RT-PCR positiven cDNAs zeigten.

Die Reaktivität von Patientenseren mit diesen Klonen lag im Durchschnitt bei etwa 29%, wobei auch zwei Extreme zu beobachten waren: Die Reaktivität gegen den Klon se14-3 zeigte sich in 11% (1/9) und gegen den Klon se1-1 in 50% (5/10) CTCL-Seren. Zwei Kontrollseren (n=10) reagierten mit Klon se20-6, der zu Klon

se20-4 homolog ist. Für se20-6 konnte gezeigt werden, daß dieser für ein unterschiedliches Peptid (72 aa) im ersten Leserahmen kodiert, das weder in se20-4 noch seinem homologen Gen HRIHFB2216 vorhanden war. Die Sequenzanalyse dieser Klone und ein Vergleich mit den homologen Gegenstücken offenbarte in einigen Fällen Insertionen, Deletionen oder Elongationen.

Es muß betont werden, daß alle getesteten Klone serologisch spezifisch sind.

Beispiel 6: Sequenzanalysen hinsichtlich Lg1-2

Es konnten eine Reihe weiterer Klone mit großer Übereinstimmung zu Lg1-2 isoliert werden, die in 3'-Richtung komplettiert waren (Stop-Codon und 3'-untranslatierte Region vorhanden). Dies ist in Fig. 24 dargestellt. Diese Klone ließen sich in einem Gen (GBP-TA) zusammenfassen, das sich dem Chromosom 1p22.3 zuordnen ließ. Dabei konnten zwei Splicing-Varianten unterschieden werden: GBP-TA wurden 12 Exons zugeordnet, während bei GBP-TA_{short} Exon Nummer 2 fehlte.

Aus GBP-TA ließ sich ein Protein ableiten, das gewisse Homologien zu den bekannten Guanylat-bindenden Proteinen GBP-1, GBP-2 und HGBP (US-A-5,871,965) aufweist. Die Sequenz von HGBP beinhaltet aber nicht Exon 2.

Beispiel 7: Expressionsanalysen hinsichtlich GBP-TA

Zur Analyse der Expression von GBP-TA wurden RT-PCR Experimente durchgeführt. Es wurden zwei verschiedene Primerpaare (s. Fig. 24) zur Unterscheidung der beiden Splicing-Varianten verwendet. Dabei wurde eine große Zahl von Kontroll-cDNAs verwendet, die jeweils aus einer Sammlung von Geweben verschiedener Spender hergestellt wurden. Während 11 Kontroll-Gewebe für beide Primer negativ waren, wurden in 5 Geweben GBP-TA_{short}, und in nur 2 Geweben (Knochenmark und Magen) beide Varianten von GBP-TA nachgewiesen (Tabelle 6).

Tabelle 6: Nachweis von GBP-TA und GBP-TA_{short} in adulten Kontrollgeweben

Kontrollgewebe	Ergebnis	
	Primerset I	PrimersetII
Hirn, Colon, Herz, Niere, Leber, Ovar, Lunge, PMNC, Prostata Testis, Thymus, Trachea	-	-
Plazenta, Dünndarm, Milz, aktiv. CD8 T-Zellen, Uterus	-	+
Knochenmark, Magen	+	+

Im Gegensatz zu den Kontrollgeweben wurde GBP-TA in verschiedenen Tumorgeweben mittels RT-PCR nachgewiesen: Kutane T-Zell-Lymphome (26%, n=19), Tumore aus dem HNO-Bereich (21%, n=14). GBP-TA_{short} wurde in 20% der HNO-Tumore (n=15) und 9% der Kolon-Karzinome (n=35) gefunden. Se57-1 wurde in 20% der Colonkarzinome (n=35) und 57% der HNO-Tumore (n=28) nachgewiesen.

Da die RT-PCR äußerst sensitiv ist und keine Aussage über das Vorhandensein und die Menge an Protein zulässt, wurde die Expression mittels Western Blot und eines GBP-TA spezifischen Antikörpers überprüft. Von den RT-PCR positiven Kontrollen konnten mehrere als Protein-Medleys im Western Blot getestet werden: Plazenta, Dünndarm, Milz, fötale Leber, Magen, Testis, Uterus. Diese erwiesen sich ebenso wie andere getestete Kontrollprotein-Medleys (Brustdrüse, Testis) als negativ, während Proteine gewonnen aus Tumor (CTLG)-Zelllinien (SeAx, MyLa, Hut-78, HH; Isolierung über Tristar, AGS, Heidelberg) eine deutliche Bande in der entsprechenden Größe zeigten. Dies

beweist eine Eignung von GBP-TA als spezifische Zielstruktur zur Therapie.

Beispiel 8: Herstellung von Antikörpern gegen GBP-TA

Zur Herstellung eines GBP-TA-spezifischen Antikörpers wurde das Insert eines Klons, das die Basen 539 bis einschließlich 1991 von GBP-TA umfaßt, in einen His-Vektor kloniert und in E-coli exprimiert. Das rekombinante Protein wurde über eine Nickel-Säule aufgereinigt, anschließend zur weiteren Reinigung in einem SDS-Gel aufgetrennt und die entsprechende Bande ausgeschnitten. Mit dem ausgeschnittenen Gelstück wurde ein Kaninchen immunisiert, dessen Präimmunserum nicht mit dem Antigen reagierte.

Immunisierungsprotokoll für polyklonale Antikörper im Kaninchen

Pro Immunisierung werden 600 µg gereinigtes KLH-gekoppeltes Peptid in 0,7 ml PBS und 0,7 ml komplettem bzw. inkomplettem Freund's Adjuvans eingesetzt.

Tag 0:	1. Immunisierung (komplettes Freund's Adjuvans)
Tag 14:	2. Immunisierung (inkomplettes Freund's Adjuvans; icFA)
Tag 28:	3. Immunisierung (icFA)
Tag 56:	4. Immunisierung (icFA)
Tag 80:	Ausbluten

Das Serum des Kaninchens wird im Immunoblot getestet. Hierzu wird das zur Immunisierung eingesetzte Peptid einer SDS-Polyacrylamid-Gelelektrophorese unterzogen und auf ein Nitrocellulosefilter übertragen (vgl. Khyse-Andersen, J., J. Biochem. Biophys. Meth. 10, (1984), 203-209). Die Western Blot-Analyse wurde wie in Bock, C.-T. et al., Virus Genes 8, (1994), 215-229, beschrieben, durchgeführt. Hierzu wird das Nitrocellulosefilter eine Stunde bei 37°C mit einem ersten Antikörper inkubiert.

Dieser Antikörper ist das Serum des Kaninchens (1:10000 in PBS). Nach mehreren Waschschritten mit PBS wird das Nitrocellulosefilter mit einem zweiten Antikörper inkubiert. Dieser Antikörper ist ein mit alkalischer Phosphatase gekoppelter monoklonaler Ziege Anti-Kaninchen-IgG-Antikörper (Dianova) (1:5000) in PBS. Nach 30-minütiger Inkubation bei 37°C folgen mehrere Waschschrritte mit PBS und anschließend die alkalische Phosphatase-Nachweisreaktion mit Entwicklerlösung (36µM 5'-Bromo-4-chloro-3-indolylphosphat, 400µM Nitroblau-tetrazolium, 100mM Tris-HCl, pH 9.5, 100 mM NaCl, 5 mM MgCl₂) bei Raumtemperatur, bis Banden sichtbar werden.

Es zeigt sich, daß erfindungsgemäße, polyklonale Antikörper hergestellt werden können.

Beispiel 9: ELISA

Das gleiche GBP-TA Insert, das zur Antikörpergewinnung verwendet wurde (Basen-Nr. 539 bis einschl. 1991 von GBP-TA) wurde in einen pGEX-Vektor kloniert, rekombinant exprimiert und in einem GST-ELISA eingesetzt. Das verwendete ELISA-System lehnt sich an Sehr et al. (J. of Immunol. Meth. 2001, 253, 153-162) an. Hierfür wurden Glutathion-Casein beschichtete ELISA-Platten mit dem Fusionsprotein beladen und anschließend Seren von CTCL Patienten und Kontrollpersonen auf das Vorhandensein von spezifischen Antikörpern getestet. Dabei zeigte sich, daß 17% der CTCL-Seren (n=60), aber nur 2% der Kontrollseren (n=99) mit dem Fusionsprotein GST-GBP-TA reagierten.

Der ELISA eignet sich für alle angegebenen Markerantigene für diagnostische Zwecke, zur Prognose-Abschätzung und zur Verlaufskontrolle.

Patentansprüche

1. Diagnostische Zusammensetzung, die mindestens eine Nukleinsäuresequenz enthält, deren veränderte Expression mit einer Tumor-Erkrankung assoziiert ist, wobei die Nukleinsäuresequenz se2-5 (Fig.1), se20-10 (Fig.2), se57-1 (Fig.3), se70-2 (Fig.4), Lg1-2 (Fig. 5), sel-1 (Fig. 6), se2-1 (Fig.7), se2-2 (Fig.8), sel4-3 (Fig.9), se20-4 (Fig.10), se20-7 (Fig.11), se20-9 (Fig.12), se33-1 (Fig.13), se37-2 (Fig.14), se89-1 (Fig.15), L14-2 (Fig.16), L15-7 (Fig.17), Li9-1 (Fig.18), Li9-4 (Fig.19), Lii5-2 (Fig. 20), Lii10-6 (Fig. 21), Liii4-5 (Fig. 22) oder GPB-TA (Fig. 23) umfaßt.

2. Arzneimittel, das mindestens eine Nukleinsäuresequenz enthält, deren veränderte Expression mit einer Tumor-Erkrankung assoziiert ist, wobei die Nukleinsäuresequenz se20-10 (Fig.2), se57-1 (Fig.3), Lg1-2 (Fig. 5), sel-1 (Fig. 6), se2-1 (Fig.7), se2-2 (Fig.8), sel4-3 (Fig.9), se20-7 (Fig.11), se20-9 (Fig.12), se33-1 (Fig.13), se37-2 (Fig.14), L14-2 (Fig.16), L15-7 (Fig.17), Li9-1 (Fig.18), Li9-4 (Fig.19), Lii5-2 (Fig. 20), Lii10-6 (Fig. 21), Liii4-5 (Fig. 22) oder GPB-TA (Fig. 23) umfaßt.

3. Diagnostische Zusammensetzung nach Anspruch 1 oder Arzneimittel nach Anspruch 2, wobei die mindestens eine Nukleinsäuresequenz, deren veränderte Expression mit einer Tumor-Erkrankung in Zusammenhang steht, eine Nukleinsäuresequenz umfaßt,

(a) die sich von einer in Anspruch 1 oder 2 definierten Nukleinsäuresequenz in der Codonsequenz aufgrund der Degeneration des genetischen Codes unterscheidet;

(b) die mit einer in einem der Ansprüche 1, 2 oder 3(a) definierten Nukleinsäuresequenz hybridisiert; oder

(c) die ein Fragment, eine allelische Variante oder eine andere Variante einer in einem der Ansprüche 1, 2, 3(a) oder 3(b) definierten Nukleinsäuresequenz ist.

4. Nukleinsäuresequenz entsprechend der Definition nach einem der Ansprüche 1 bis 3, die eine cDNA oder genomische DNA ist.
5. Protein, dessen veränderte Konzentration mit einer Tumorerkrankung in Zusammenhang steht und das von einer Nukleinsäuresequenz nach einem der Ansprüche 1 bis 4 kodiert ist.
6. Diagnostische Zusammensetzung, die mindestens einen Vektor enthaltend eine der Nukleinsäuresequenzen nach Anspruch 1 oder 3, mindestens ein von einer Nukleinsäuresequenz nach Anspruch 1 oder 3 kodiertes Protein oder mindestens einen gegen dieses Protein gerichteten Antikörper enthält.
7. Diagnostische Zusammensetzung nach Anspruch 6 zur Diagnose oder Verlaufskontrolle einer Tumorerkrankung.
8. Diagnostische Zusammensetzung nach Anspruch 7, wobei diese in Form eines ELISA, Protein-Chips, Nukleinsäure-Chips oder einer mit DNA, RNA oder Protein beladener Membran bereitgestellt wird.
9. Arzneimittel, das mindestens einen Vektor enthaltend eine der Nukleinsäuresequenzen nach Anspruch 2 oder 3, mindestens ein von einer Nukleinsäuresequenz nach Anspruch 2 oder 3 kodiertes Protein oder mindestens einen gegen dieses Protein gerichteten Antikörper enthält.
10. Arzneimittel nach Anspruch 9 zur Therapie von Tumorerkrankungen.
11. Verwendung einer Nukleinsäuresequenz nach einem der Ansprüche 1 bis 4 zur Diagnose und/oder Therapie einer Tumorerkrankung.

12. Verwendung mindestens eines von einer Nukleinsäuresequenz nach einem der Ansprüche 1 bis 4 kodierten Proteins zur Diagnose und/oder Therapie einer Tumorerkrankung.

13. Verwendung mindestens eines gegen ein von einer Nukleinsäuresequenz nach einem der Ansprüche 1 bis 4 kodierten Proteins gerichteten Antikörpers zur Diagnose und/oder Therapie einer Tumorerkrankung.

14. Verwendung nach Anspruch 12 oder 13 als Vakzinierungsmittel.

15. Verwendung nach Anspruch 12 zur Erzeugung Peptid-beladener Antigen-präsentierender Zellen (APC).

16. Verwendung nach Anspruch 12 zur Erzeugung tumorspezifischer T- Zellen.

17. Diagnostische Zusammensetzung nach Anspruch 1, 3 oder 6, Arzneimittel nach Anspruch 2, 3 oder 9, Verwendung nach einem der Ansprüche 11 bis 17, wobei der Tumor CTCL ist.

se2-5

2234 bp

1/44

BASE COUNT 711 a 470 c 566 g 487 t

ORIGIN

```
1 aggatgagga tgggacagaa gaggataaca gtcgtgttga acctgttgga catgctgaca
61 cgggttttga gcatataccc aactttttctc tggatgatat ggtaaagctc gtagaagtcc
121 ccaacgatgg agggcctctg ggaatccatg tagtgccttt cagtgcctga ggcggcagaa
181 ccctgggggtt attagtaaaa cgattggaga aagggtgtaa agctgaacat ctgaacatga
241 aaatcttttt cgtgagaatg attgcattgt caggattaat gatggcgacc ttcgaaatag
301 aagatttgaa caagcacaac atatgtttcg ccaagccatg cgtacacca tcatattggtt
361 ccatgtgggtt cctgcagcaa ataaagagca gtatgaacaa ctatcccaaa gtgagaagaa
421 caattactat tcaagccgtt ttagccctga cagccagtat attgacaaca ggagtgtgaa
481 cagtgcaggg cttcacacgg tgcagagagc accccgactg aaccaccgcg ctgagcagat
541 agactctcac tcaagactac ctcatagcgc acaccctcg ggaaaaccac catccgctcc
601 agcctcggca cctcagaatg tatttagtac gactgtaagc agtggttata acaccaaaaa
661 aataggcaag aggcttaata tccagcttaa gaaaggtaca gaaggtttgg aattcagcat
721 cacttccaga gatgtaacaa taggtggctc agctccaatc tatgtgaaaa acattctccc
781 cgggggggcg gccattcagg atggccgact taaggcagga gacagactta tagaggtaaa
841 tggagtagat ttagtgggca aatcccaaga ggaagttgtt tcgctgttga gaagcaccaa
901 gatggaagga actgtgagcc ttctggtctt tcgccaggaa gacgccttcc acccaaggga
961 actgaatgca gagccaagcc agatgcagat tccaaaagaa acgaaagcag aagatgagga
1021 tattgttctt acacctgatg gcaccaggga atttctgaca ttggaagtcc cacttaatga
1081 ttcaggatct gcaggccttg gtgtcagtgt caaaggtaac cggtaaaaag agaaccacgc
1141 agatttgga atctttgtca agtccattat taatggagga gcagcatcta aagatggaag
1201 gcttcgggtg aatgatcaac tgatagcagt aaatggagaa tccctgttgg gcaagacaaa
1261 ccaagatgcc atggaaaccc taagaaggtc tatgtctact gaaggcaata aacgaggaat
1321 gatccagctt attgttgcaa ggagaataag caagtgcaat gagctgaagt cacctgggag
1381 ccccccctga cctgagctgc ccattgaaac agcgttggat gatagagaac gaagaatttc
1441 ccattccctc tacagtggga ttgaggggct tgatgaatcg cccagcagaa atgctgccct
1501 cagtaggata atgggtgagt caggtaaata ccagctgtcc cctacagtga atatgcccc
1561 agatgacact gtcattatag aagatgacag gttgccagtg cttcctccac atctctctga
1621 ccagtcctct tccagctccc atgatgatgt ggggtttgtg acggcagatg ctggtacttg
1681 ggccaaggct gcaatcagtg attcagccga ctgctctttg agtccagatg ttgatccagt
1741 tcttgctttt caacgagaag gatttggacg tcagatagct gacgagacta aactcaatac
1801 agtggatgac cagaaagcag gttctcccag cagagatgtg ggtccttccc tgggtctgaa
1861 gaagtcaagc tcgttgagga gtctgcagac cgcagttgcc gaggtgactt tgaatgggga
1921 tattcctttc catcgtccac ggccgcggat aatcagaggc aggggatgca atgagagctt
1981 cagagctgcc atcgacaaat cttatgataa acccgcggtg gatgatgatg atgaaggcat
2041 ggagaccttg gaagaagaca cagaagaaag ttcaagatca gggagagagt ctgtatccac
2101 agccagtgat cagccttccc actctctgga gagacaaatg aatggaaacc aagagaaagg
2161 tgataagact gatagaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa
2221 aaaaaaaaaa aaaa
```

Fig. 1 (1)

se2-5 protein ⁶⁵⁴aa561

1 mfrqamrtpi iwfhvvpaaan keqyeqlsqe eknnyyssrf spdsqyidnr svnsaglhv
61 qraprlnhpp eqidshsrhp hsahpsgkpp sapasapqnv fsttvssgyn tkkigkrlni
121 qlkkgtogle fsitsrdvti ggsapiyvkn ilprgaaiqd grlkagdrli evngvdlvgk
181 sqeevvsllr stkmegtvs1 lvfrqedafh prelnaepsq mqipketkae dedivltpdg
241 trefltfevp lndsgsaglg vsvkgnrske nhadlgifvk siinggaask dgrlrvndql
301 iavngesllg ktnqdametl rrmstegnk rgmiqlivar riskcnelks pgsppgpelp
361 ietalddrer rishslysgi egldespsrn aalsrimges gkyqlsptvn mpqddtviie
421 ddrlpvlpph lsdqsssssh ddvgfvtada gtwakaaaisd sadcslspdv dpvlafqreg
481 fgrqiadetk lntvddqkag spsrdvgpsl glkksssles lqtavaevtl ngdipfhrpr
541 priirgrgcn esfraaidks ydkpavdddd egmetleedt eessrsgres vstasdqpsh
601 slerqmngnq ekgdktdrkk kkkkkkkkkk kkkkkkkk

//

Fig. 1 (2)

se20-10 1991 bp

3/44

BASE COUNT 776 a 360 c 392 g 463 t

ORIGIN

1 aaaacgcttt ttgcatacaa gcaggaaaat gagatgttat ccagtagtag agatcagaga
61 gttgtgacat ctgaggacca agttcaagaa gggactaaag tgctgaaact taaaacaaaa
121 atggctgata aagaaaacat gaagagacct gcagagagca aaaataatac agtgggtggg
181 aaacattgta ttcctttaaa accttcaaat gaactaacca attcaactgt agtaattgac
241 acacataaac ctaaggatag taatcaaact ccgcatttgt tactaactga agatgatccc
301 caaagtcaac atatgacatt aagccaggca ttccacctta aaaacaatag taaaaagaaa
361 caaatgacta cagaaaaaca aaagcaagat gctaacaatgc ccaagaaacc tgtgcttgga
421 tcttatcgtg gccagattgt tcagtctaag attaattcat ttagaaaacc tctacaagtc
481 aaagatgaga gttctgcagc aacaaagaaa ctttcagcca ctatacctaa agccacaaaa
541 cctcagcctg taaacaccag cagtgtaca gtgaaaagta atagatcctc caataagact
601 gccactacta aatttgtgag cactacatct cagaacacac aacttgtgcg acctcctatt
661 agaagtcac acagtaatac ccgggacact gtgaaacaag gcatcagtag aacctctgcc
721 aatgttaca tccggaaagg gcctcatgaa aaagaactat tacaatcaaa aacagcttta
781 tctagtgtca aaaccagttc ttctcaaggt ataataagaa ataagactct atcaagatcc
841 atagcatctg aagttgtagc caggcctgct tcattgtcta atgataaact gatggaaaag
901 tcagagcccg ttgaccagcg aagacatact gcaggaaaag caattgttga tagtagatca
961 gctcagccca aagaaacctc ggaagagaga aaagctcgtc tgagtgtgtg gaaagctggc
1021 aaaggaagag tgctaaaaag gccccctaatt tcagtagtta ctcagcatga gcctgcagga
1081 caaatgaaa aaccagttgg gtcttttttg actaccatgg cagaagaaga tgaacaaaga
1141 ttatttactg aaaaagtaaa caacacattt tctgaatgcc tgaacttgat taatgaggga
1201 tgtccaaaag aagatatact ggtcacactg aatgacctga ttaaaaaatat tccagatgcc
1261 aaaaagcttg ttaagtattg gatatgtctt gcacttattg aaccaatcac aagtcctatt
1321 gaaaatatta ttgcaatcta tgagaaagcc attctggcag gggctcagcc tattgaagag
1381 atgcgacaca cgattgtaga tattctaaca atgaagagtc aagaaaaagc taatttagga
1441 gaaaatatgg agaagtcttg tgcaagcaag gaagaagtca aagaagtcag tattgaagat
1501 acaggtgttg atgtagatcc agaaaaactg gaaatggaga gtaaaacttca tagaaatttg
1561 ctattttcaag attgtgaaaa agagcaagac aacaaaacaa aagatccaac ccatgatgtt
1621 aaaacccccca atacagaaac gaggacaagt tgcttaatta aatataatgt gtctactacg
1681 ccatacttgc aaagtgtgaa aaaaaaaggc gcagtttgat ggaacaaatt ccgcatttaa
1741 agagctgaag tttttaacac cagtgaagac ttctcgacgt cttcaagaga aaacttctaa
1801 attgccagat atgttaaaag atcattatcc ttgtgtgtct tcattggaac agctaacgga
1861 gttgggaaga gaaactgatg cttttgtatg ccgccctaatt gcagcactgt gccgggtgta
1921 ctatgaggct gataacaacat aagagaaata aagctctgtt agggaaaaaa aaaaaaaaaa
1981 aaaaaaaaaa a

//

Fig. 2 (1)

se20-10 protein 5⁴ aa

4/44

1 mlsssrldgrv vtseqdvqgeg tkvlklktkm adkenmkrpa esknntvvgk hciplkpsne
61 ltnstvvidt hkpkdsngtp hlllteddpq sqhmtlsqaf hlknskkkq mtekkqkqda
121 nmpkpkpvlgs yrgqivqski nsfrkplqvkd dessaatkkl satipkatkp qpvtssvtv
181 ksnrssnkta ttkfvsttsq ntqlvrppir shhsnttrdtv kqgisrtsan vtirkgphek
241 ellqsktals svktsssqgi irnktlsrsi asevvvarpas lsndklmeks epvdqrrhta
301 gkaivdsrsa qpketseerk arlsewkagk grvlkrppns vvtqhepagq nekpvgswft
361 tmaeedeqrl ftekvnttfs eclnlinegc pkedilvtln dliknipdak klvkywicla
421 liepitspie niiaiyekai lagagpieem rhtivdiltm ksqekanlge nmekscaske
481 evkevsiedt gvdvdpekle mesklhrnll fgdcekeqdn ktkdpthdvk tpntetrts
541 likynvsttp ylvsvkkkga v

//

Fig. 2 (2)

ie57-1 3997 bp

BASE COUNT 1375 a 662 c 743 g 1217 t

ORIGIN

```
1 ctggcccaaa ggtccgatcg cccaggggag gagcagcacc gggaccccgc gtcggctggg
61 cgccccacaa gggaagccag tcttaatatg atggaaacat ctctgaactt ctaaaagacc
121 aagggttgcg ttttagctct attaatTTTA cttcgtcttg gccagaattc acaatgacaa
181 cagtgcacgt gaccacagaa attcccccaa gggataagat ggaagataat tctgccttgt
241 atgagtctac gtccgctcac attattgaag aaaccgagta tgtgaaaaag attcgaacta
301 ctctgcaaaa gatcaggacc cagatgttta aagatgaaat aagacatgac agtacaaatc
361 acaaactaga tgcaaagcac tgtggaaacc ttcaacaggg ctctgattct gaaatggatc
421 cttcttgttg cagtttggat ttgcttatga aaaagataaa aggaaaagac ctacagctct
481 tagaaatgaa caaagagaat gaagtattga aaatcaagct gcaagcctcc agagaagcag
541 gagcagcagc tctgagaaac gtggcccaga gattatttga aaactaccaa acgcaatctg
601 aagaagtgag aaagaagcag gaggacagta aacaattact ccaggttaac aagcttgaaa
661 aagaacagaa attgaaacaa catgttgaaa atctgaatca agttgctgaa aaacttgaag
721 aaaaacacag tcaaattaca gaattggaga accttgtaaa gagaatggaa aaggaaaaga
781 gaacactact agaaagaaaa ctgtctttgg aaaacaagct actgcaactc aaatccagtg
841 ctacatatgg aaaaagttgc caggatcttc agagggagat ttccattctc caggagcaga
901 tctctcatct gcagtttgtg attcaactcc aacatcagaa cctgcgcagt gtcattccagg
961 agatggaagg attaaaaaat aattttaaag aacaagacaa aagaattgaa aatctcagag
1021 aaaagggttaa catacttgaa gccagaata aagaactaaa aaccaggta gcactttcat
1081 ctgaaactcc taggacaaaag gtatctaagg ctgtctctac aagtgaattg aagaccgaag
1141 gtgtttcccc ttatttaatg ttgattaggt tacggaaatg aactggctgg atgaagatct
1201 gatttagaaa gactgcgtga gtcttattta ttctctgaaa cacagcccaa gtttcatgtt
1261 aaaatggcaa aatgccatta tttaaatgga acttattaca taccaatggc tttgcaagaa
1321 gatgacattt cagaagatca aacaaatcta tatttaatgg atggactctt caaaacttac
1381 caaatagttg aagaaaccag gtgccttctc atgatggaag acagattctg ctttaaatta
1441 aaaaaaaaaa aaatctgaat cttgttttca gatttttttt tctactggga ttgttttaag
1501 attgtcaatt ctgacttttt tatagtgggt ttttaagagta taaatagaag ggagagtgtg
1561 tatgtgtatg aatgaacata catttctctg atatatatgt atgaagggca tgtatatgta
1621 tgaatgagca aacatatttg aaagttaact tttggatgat aggaaagatc gtacagtgc
1681 ataagttcat ctctgatcc attgtttgtg ggagaattat acttgactga attatgggca
1741 ggagaaagag cagattcctt ttagctaatt cccaaccca tatgcccctc tgaagttgag
1801 aatcatggct gcctcaccac acatcagaga atgactgttc ttcttagttc tgggattaaa
1861 aattggtttc tagaggtaac ctgtacacac aaacgcacag ggatgcacat ggtttctctc
1921 gcctttgtga ctaattttct tcttgatagt tattaatagt atctaaataa aatattgggg
1981 ggatagaaaa ataatgctgt tagctcatac ttccatgaaa atgtatatat tataggctca
2041 aaggaataat gactgctgtc tgcagccaga aagaatctga atttatgaat tggaaagata
2101 tatatagtgt atttgtgaaa gtttgcttaa attctgatac atgccttctt tgtaggtgga
2161 gtttgtgatt gcagtgaata gaacaaattc tgacttgagc aaatgcagac atacagtcag
2221 gaaagaacaa acttcaatta aatgtataat gagagacctt ggtcccccta aaggatgaat
```

6/44

2281 ttcccttttagg cctttgatct tcctctctag tgtataactt taaat²attg ctcagaaaag
2341 atgctgactc ttccattatg gaatgtgaaa tatcagtgtt gtctataaat atttgaaggt
2401 atataaaaaat gagatgatgt aatgtattta taaatttatc caagtactgt aatccttgaa
2461 ttgttgtgta actgtgtgtg agttttatgc ttcattggtat ttttggaac atttttattg
2521 cttgcttatt ttgaaggtat atttatctgt taacatttag ggcatagtta ccttatacca
2581 gcaacaagtc taagcacttt acttgattta atttatttta tcttcccaa ggccctcaga
2641 agcagattat tcctgtttta cagagaaaga tactgagagt ggttgagtaa tttgcccagc
2701 tgctaactgt gaagcaaaaa tttgaacctt ttgacttggc tattgatatt cattctactt
2761 gctcacatgg tggcttaaga aatttcccag ctatagaaat ctctctattt ttgccacttt
2821 aatcaacaca tagcttcctg gatgactgcc tgtgttattt tgtggatgac agtaagaaac
2881 aacaaatact gataaaatca atattttgct gaaatgagtt gatctttcac cagctggact
2941 accatttgtga gaactcagtt cagacaaact tccctgctaa aaatctgttt atcatacatt
3001 tattatttat gactttatgt cacattgaag aatttcttca tgatacattt tcaggcacac
3061 ttgtaggaaa attaggatca tgagtctgc ttttaagtatt tgcagtgtag taagagaatc
3121 catcttttac taggagacca gattcctttt atacctcatt catcatgctg gattgtaata
3181 aatttcagat tttggaatgg gcttatttaa ctgacctaac aatcttgatg atttccatta
3241 gaataactta ttctaaggtc aaaagtggaa agacactgtt ggtttttatt ttgatttcac
3301 tatactcatt tttgaacatg gaaatacagt ggtgaaacca cctatgcaaa aatgataaca
3361 gtgaggaaat tatgacagtg aaagagatct gacctaaacta tctatcttgc ctcgaaactg
3421 cccttggtcg ttctgagtg tgggccaaag taactttggg agaaatttac tttatagggt
3481 aaattataat agcccttccc aaaactaac gattctcctg cctcagcctc ccgagtagct
3541 gtctttataa taccatcagc cttatcattt attcgtcatg tatggattgt ttcctatatac
3601 cactatcata aaattatcat ttgaaatatt tttttatgaa aaaaaaaca cttctcagtg
3661 aaaaaacaag attacaaggg ggaataatat ttattcagct aaaatagtag ttcacaggaa
3721 atataggaaa agaaacacta gttcagtttt attccaaata atgtacttct aattatactc
3781 ttgaatttat tggataagaa ggtctgaggt gggcttgaaa ttcatacatg aagtcaggaa
3841 aagaaaagaa acttagttct accttgatta catatgttgt taaaattat cattataaaa
3901 tgttttaaca attagtatag tatatctttt gaataattgc ttataatatg ccttaccata
3961 aagaaaattg atgctaaaaa aaaaaaaaaa aaaaaaa

//

se57-1 protein 335aa

1 mttvtvttei pprdkmedns alyestsahi ieeteyvkki rttlqkirtg mfkdeirhds
61 tnhkldakhc gnlqggsdse mdpsccsldl lmkkikgkdl qllemnkene vlkiklqasr
121 eagaaalrnv aqrlfenyqt qseevrkkqe dskqllqvnk lekeqklkqh venlnqvaek
181 leekhsqite lenlvqrmek ekrtllerk lslenkllqk ssatygkscq dlqreisilq
241 eqishlqfvi hsqhqnlrsv iqemeglknn lkeqdkrien lrekvnilea qnkelktgva
301 lissetprtkv skavstselk tegvspyml irlrk

//

Fig. 3 (2)

7/44

se70-2 1592 bp

BASE COUNT 577 a 230 c 335 g 450 t

ORIGIN

```

1 ctgctttgaa ggctgcacag aaaaccttac ttgtttccac ctctgcagtt gataataatg
61 aagcacagaa aaaaaaacag gaggcattga aacttcagca ggatgtaagg aaaaggaaac
121 aagaaatttt agaaaagcac attgaaacac agaagatggt aatttcaaaa ctggagaaaa
181 acaaaacaat gaagtctgaa gataaagcag aaataatgaa aacttttagag gttttgacaa
241 aaaatattac caagttgaaa gatgaggtca aagctgcttc tcctggacgc tgtcttccaa
301 aaagtataaa aaccaagact cagatgcaga aggaattact tgacacagaa ctggatttat
361 ataagaagat gcaggctgga gaagaagtca ctgaacttag gagaaagtat acagaattac
421 agctggaagc tgccaaacga gggattcttt catctggtcg gggcagagga attcattcaa
481 gaggtcgagg tgcagttcat ggccgaggca gggggcgagg gcgagggcga ggtgtgcctg
541 gtcatgctgt ggtggatcac cgtcccaggg cattggagat ttctgcattt acggagagcg
601 atagagaaga tcttcttcct cattttgctc aatatggtga aattgaagat tgtcagattg
661 atgattcctc acttcatgca gtaattacat tcaagacaag agcagaagct gaagcagctg
721 cagttcatgg agctcgtttc aaagggaag atctaaaact ggcatggaat aaaccagtaa
781 ctaatatttc agctgttgaa acagaagaag ttgagcctga tgaagaagaa tttcaggaag
841 agtctttggt ggatgactca ttacttcaag atgatgatga agaagaagag gacaatgaat
901 ctctgttctg gagaagatga ttgactgat cattgatctg catatgctag aactctacct
961 gtgtttcatt agtattatct aatgtacttt tacatatttg taaaaacaat ttttggtaaa
1021 atgtgatgaa gatggatttc acaaatagac aaaaaagaag aaaactacct tctgatcttg
1081 tattttgaaa gattgatggt tgcattttac ttcagtaaac aattgctaaa gacatcacac
1141 tagaaacata tgcaatgttt ttattacata cttctactgg acatcacaga attctttggg
1201 ttctttgtaa tttaatgaat aggtctgaaa acttatgacc aatacttggt ataacttaga
1261 ggactttggt ttattccaaa taaggaatga atttgcattt aaaatcttaa tgaatgtttt
1321 caaaactgaa tagataacat agtactctaa ctaaagtctc caagttatgt attataatat
1381 tacatagtag tatgcttagg ctttactatg tattagcctt ttgttggaact gtgtatgtat
1441 tttaccatat gggttttaat gataatggtg tatgactgct ttacatgagt ccttatgcat
1501 ccagatgtta taataaagtg gaatggtctc tttaaaaaaa aaaaaggaaa gaaaagagaa
1561 aagcaatgac aaaaaaaaaa aaaaaaaaaa aa

```

//

se70-2 protein 254aa

```

1 mliskleknk tmksedkaei mktlevltkn itklkdevka aspgrclpks iktktqmcke
61 lldteldlyk kmqageevte lrrkytelql eaakrgilss grgrgihsrg rgavhgrgrg
121 rgrgrgvpgh avvdhrpral eisaftesdr edllphfaqy geiedcqidd sslhavitfk
181 traeeaaav hgarfkqgdl klawnkpvt n isavetevee pdeefqees lvddslldd
241 deeednesr swrr

```

//

Fig. 4

g1-2

BASE COUNT 600 a 478 c 446 g 456 t

ORIGIN

```
1 aattcggcac gaggttcact ctttgcaata aatcttgctg ctgctcactc tttgggtcca
61 cactgccttt atgagctgta acactcactg ggaatgtctg cagcttcact cctgaagcca
121 gcgagaccac gaacccacca ggaggaacaa acaactccag acgcgcagcc ttaagagctg
181 taacactcac cgcgaaggtc tgcagcttca ctctgagcc agccagacca cgaacccacc
241 agaaggaaga aactccaaac acatccgaac atcagaagga gcaaactcct gacacgccac
301 ctttaagaac cgtgacactc aacgctaggg tccgcggctt cattcttgaa gtcagtgaga
361 ccaagaaccc accaattccg gacacgctaa ttgttgtaga tcatcacttc aagggtgcca
421 tatctttcta gtggaaaaat tattctggcc tccgctgcat acaaactcagg caaccagaat
481 tctacatata taaggcaaag taacatccta gacatggctt tagagatcca catgtcagac
541 cccatgtgcc tcatcgagaa ctttaatgag cagctgaagg ttaatcagga agctttggag
601 atcctgtctg ccattacgca acctgtagtt gtggtagcga ttgtgggcct ctatcgact
661 ggcaaatcct acctgatgaa caagctggct gggaagaaca agggcttctc tgtgcatcta
721 cgggtgcagtc tcacaccaag ggaatttggg tatggtgtgt gcctcatccc aactggccaa
781 atcacacatt agttctgctt gacaccgagg gcctgggaga tgtagagaag gctgacaaca
841 agaatgatat ccagatcttt gcaactggcac tcttagtgag cagcaccttt gtgtacaata
901 ctgtgaacaa aattgatcag ggtgctatcg acctactgca caatgtgaca gaactgacag
961 atctgctcaa ggcaagaaac tcacccgacc ttgacagggt tgaagatcct gctgactctg
1021 cgagcttctt cccagactta gtgtggactc tgagagattt ctgcttaggc ctggaaatag
1081 atgggcaact tgtcacacca gatgaatacc tggagaattc cctaaggcca aagcaaggta
1141 gtgatcaaag agttcaaaat ttcaatttgc cccgtctgtg tatacagaag ttctttccaa
1201 aaaagaaatg ctttatcttt gacttacctg ctacacaaaa aaagcttgcc caacttgaaa
1261 cactgcctga tgatgagcta gagcctgaat ttgtgcaaca agtgacagaa ttctgttctt
1321 acatctttag ccattctatg accaagactc ttccagggtg catcatggtc aatggatctc
1381 gtctaaagaa cctggtgctg acctatgtca atgccatcag cagtggggat ctgccttgca
1441 tagagaatgc agtcctggcc ttggctcaga gagagaactc agctgcagtg caaaaggcca
1501 ttgcccacta tgaccagcaa atgggcccaga aagtgcagct gcccatggaa accctccagg
1561 agctgctgga cctgcacagg accagtgaga gggaggccat tgaagtcttc atgaaaaact
1621 ctttcaagga tgtagacca agtttccaga aagaattgga gactctacta gatgcaaaac
1681 agaatgacat ttgtaaacgg aacctggaag catcctcgga ttattgctcg gctttactta
1741 aggatatttt tggctctcta gaagaagcag tgaagcaggg aatttattct aagccaggag
1801 gccataatct cttcattcag aaaacagaag aactgaaggc aaagtactat cgggagcctc
1861 ggaaaggaat acaggctgaa gaagttctgc agaaatattt aaagtccaag gagtctgtga
1921 gtcatgcaat attacagact gaccaggctc tcacagagac ggaaaaaaaa aaaaaaaaaa
```

//

Fig. 5 (1)

Lgl-2 protein 214aa

9/44

1 mtktlpggim vngsrlnlv ltyvnaissg dlpkienavl alaqrnsaa vqkaiahdyd
61 qmgqkvqlpm etlgelldlh rtseraiev fmksfkdv qsfqkeletl ldakqndick
121 rnleassdyc sallkdifgp leeavkqgiy skpgghnlfi qkteelkaky yreprkgiqa
181 eevlqkylks kesvshailq tdqaltetek kkkk

//

Fig. 5 (2)

10/44

sel-1 2363 bp

BASE COUNT 1009 a 310 c 463 g 581 t

ORIGIN

1 tgaatacgca attagaactt tcagaacaac ttaaatttca gaacaactct gaagataatg
61 ttaaaaaact acaagaagag attgagaaaa ttaggccagg ctttgaggag caaattttat
121 atctgcaaaa gcaattagac gctaccactg atgaaaagaa ggaaacagtt actcaactcc
181 aaaatatcat tgaggctaatt tctcagcatt accaaaaaaaa tattaatagt ttgcaggaag
241 agctttttaca gttgaaagct atacaccaag aagaggtgaa agagttgatg tgccagattg
301 aagcatcagc taaggaacat gaagcagaga taaataagtt gaacgagcta aaagagaact
361 tagtaaaaca atgtgaggca agtgaaaaga acatccagaa gaaatatgaa tgtgagttag
421 aaaattttaag gaaagccacc tcaaatgcaa accaagacaa tcagatatgt tctattctct
481 tgcaagaaaa tacatttgta gaacaagtag taaatgaaaa agtcaaacac ttagaagata
541 ccttaaaaga acttgaatct caacacagta tcttaaaaga tgaggtaact tatatgaata
601 atcttaagtt aaaacttgaa atggatgctc aacatataaa ggatgagttt tttcatgaac
661 gggaagactt agagttttaa attaatgaat tattactagc taaagaagaa cagggctgtg
721 taattgaaaa attaaaatct gagctagcag gtttaaataa acagttttgc tatactgtag
781 aacagcataa cagagaagta cagagtctta aggaacaaca tcaaaaagaa atatcagaac
841 taaatgagac atttttgtca gattcagaaa aagaaaaatt aacattaatg tttgaaatac
901 agggctctta ggaacagtgt gaaaacctac agcaagaaaa gcaagaagca attttaaatt
961 atgagagttt acgagagatt atggaaattt tacaacaga actgggggaa tctgctggaa
1021 aaataagtca agagttcgaa tcaatgaagc aacagcaagc atctgatgtt catgaactgc
1081 agcagaagct cagaactgct tttactgaaa aagatgccct tctcgaaact gtgaatcgcc
1141 tccagggaga aaatgaaaag ttactatctc aacaagaatt ggtaccagaa cttgaaaata
1201 ccataaagaa ccttcaagaa aagaatggag tatacttact tagtctcagt caaagagata
1261 ccatgttaaa agaattagaa ggaaagataa attctcttac tgaggaaaaa gatgatttta
1321 taaataaact gaaaaattcc catgaagaaa tggataattt ccataagaaa tgtgaaaggg
1381 aagaaagatt gattcttgaa cttgggaaga aagtagagca aacaatccag tacaacagtg
1441 aactagaaca aaaggtaaatt gaattaacag gaggactaga ggagacttta aaagaaaagg
1501 atcaaaatga ccaaaaacta gaaaaactta tggttcaaat gaaagttctc tctgaagaca
1561 aagaagtatt gtcagctgaa gtgaagtctc tttatgagga aacaataaac tcagttcaga
1621 aaaaaaaaca gttgagtagg gatttgaggg tttttttgtc tcaaaaagaa gatgttatcc
1681 ttaagaaca tattactcaa ttagaaaaga aacttcagtt aatggttgaa gagcaagata
1741 atttaataaa actgcttgaa aatgagcaag ttcagaagtt atttgttaaa actcagttgt
1801 atggttttct taaagaaatg ggatcagaag tttcagaaga cagtgaagag aaagatgttg
1861 ttaatgtcct acaggcagtc ggtgaatcct tggcaaaaat aatgaggaa aatgcaacc
1921 tggcttttca gcgtgatgaa aaagtattag agttagaaaa agagattaag tgccttcaag
1981 aagagagtgt agttcagtg gaagaactta agtctttatt gagagactat gagcaagaga
2041 aagttctctt aaggaaagag ttagaagaaa tacagtcaga aaaagaggcc ctgcagtcgt
2101 atcttctaga aatgaagaat gctaataaaa aaacaaggct tgaaaatcag aatcttttaa
2161 ttcaagttga agaagtatct caaacatgta gcaaaagtga aatccataat gaaaaagaaa
2221 aatgttttat aaaggaacat gaaaacctaa agccactact agaacaaaaa gaattacgag

11/44

2281 ataggagagc agagttgata ctattaaagg attccttagc aaaatcacct tactgtaaaa
2341 aatgatacct ctgtcttcag taa

//

sel-1 protein 685aa

1 mcqieasake heaeinklne lkenlvkqce asekniqkky ecelenlrka tsnanqdnqi
61 csillqentf veqvvnkvk hledtlkele sqhsilkdev tymnnlklkl emdaqhikde
121 ffheredlef kinelllake eggcvieklk selaglnkqf cytveqhnre vqslkeqhqk
181 eiselnetfl sdsekekltl mfeiaggkeq cenlqqekqe ailnyeslre imeilqtelg
241 esagkisqef esmkqqqasd vhelqqklrt aftekdale tvnrlqgene kllsqqelvp
301 elentiknlq ekngvyllsl sqrdtmlkel egkinsltee kddfinklkn sheemdnfhk
361 kcereerlil elgkkveqti qynselegkv neltggleet lkekdnqdk leklmvqmkv
421 lsedkevlsa evkslyeeti nsvqkkkqls rdlevflsqk edvilkehit qlekkqlmv
481 eeqdnlklk eneqvqklfv ktqlygfike mgsevsedse ekdvvnvlqa vgeslakine
541 ekcnlafqrd ekvlelekei kclqeesvq ceelksllrd yeqekvllrk eleeiqseke
601 alqsdllmk nanektrlen qnlliqveev sqtcskseih nekekcfike henlklpleq
661 kelrdrrael illkdslaks pyckk

//

Fig. 6 (2)

12/44

se2-1 2564 bp

BASE COUNT 1111 a 368 c 524 g 561 t

ORIGIN

1 ccacgcgctc cgggccgctc aggctgagcg atttcccgcc ttttctgagg ttctgaggcg
61 ggagccattg gttctttctg ttgccctcat agaccgtatg tagcagttcg cgtgggcaca
121 gaacccacgg tttcccgtca gttcttcaaa gtagatattt acaaccgtaa cagagaaaaat
181 ggaaaagcaa aagccctttg cattgttcgt accaccgaga tcaagcagca gtcagggtgc
241 tgcggtgaaa cctcagaccc tgggaggcga ttccacttcc ttcaagagtt tcaacaaatg
301 tactgaagat gattttgagt ttccatttgc aaagactaat ctctccaaaa atggggaaaa
361 cattgattca gatcctgctt tacaaaaagt taatttcttg cccgtgcttg agcagggttg
421 taattctgac tgtcactatc aggaaggact aaaagactct gatttgagga attcagaggg
481 attgagcaga gtgtattcaa aactgtataa ggaggctgaa aagataaaaa aatggaaagt
541 aagtacagaa gctgaactga gacagaaaga aagtaagttg caagaaaaca gaaagataat
601 tgaagcacag cgaagaagcca ttccaggaact gcaatttgga aatgaaaag taagtttgaa
661 attagaagaa ggaatacaag aaaataaaga ttaataaaaa gagaataatg ccacaaggca
721 tttatgtaat ctactcaaag aaacctgtgc tagatctgca gaaaagacaa agaaatatga
781 atatgaacgg gaagaaacca ggcaagttta tatggatcta aataataaca ttgagaaaat
841 gataacagct tttgaggaac ttcgtgtgca agctgagaat tccagactgg aaatgcattt
901 taagttaaag gaagattatg aaaaaatcca acaccttgaa caagaataca agaaggaaat
961 aaatgacaag gaaaagcagg tatcactact attgatccaa atcactgaga aagaaaataa
1021 aatgaaagat ttaacatttc tgctagagga atccagagat aaagttaatc aattagagga
1081 aaagacaaaa ttacagagtg aaaacttaaa acaatcaatt gagaaacagc atcatttgac
1141 taaagaacta gaagatatta aagtgtcatt acaagaaggt gtgagtactc aaaaggcttt
1201 agaggaagat ttacagatag caacaaaaac aatttgtcag ctaactgaag aaaaagaaac
1261 tcaaatggaa gaatctaata aagctagagc tgctcattcg tttgtggtta ctgaatttga
1321 aactactgtc tgcagcttgg aagaattatt gagaacagaa cagcaaagat tggaaaaaaa
1381 tgaagatcaa ttgaaaatac ttaccatgga gcttcaaaag aaatcaagtg agctggaaga
1441 gatgactaag cttacaaata acaagaaggt agaacttgaa gaattgaaaa aagtcttggg
1501 agaaaaggaa acacttttat atgaaaataa acaatttgag aagattgctg aagaattaaa
1561 aggaacagaa caagaactaa ttggtcttct ccaagccaga gagaaagaag tacatgattt
1621 ggaaatacag ttaactgcca ttaccacaag tgaacagtat tattcaaaag aggttaaaga
1681 tctaaaaact gagcttgaaa acgagaagct taagaatact gaattaactt cacactgcaa
1741 caagctttca ctagaaaaca aagagctcac acaggaaaca agtgatatga ccctagaact
1801 caagaatcag caagaagata ttaataataa caaaaagcaa gaagaaagga tgttgaaaca
1861 aatagaaaat cttcaagaaa cagaaacca attaagaaat gaactagaat atgtgagaga
1921 agagctaaaa cagaaaagag atgaagttta atgtaaattg gacaagagtg aagaaaattg
1981 taacaattta aggaacaag ttgaaaataa aaacaagtat attgaagaac ttcagcagga
2041 gaataaggcc ttgaaaaaaa aaggtacagc agaaagcaag caactgaatg tttatgagat
2101 aaaggtcaat aaattagagt tagaactaga aagtgcacaa cagaaatttg gagaaatcac
2161 agacacctat cagaaagaaa ttgaggacaa aaagatatca gaagaaaatc ttttggaaga
2221 gggttgagaaa gcaaaagtaa tagctgatga agcagtaaaa ttacagaaag aaattgataa

2281 gcgatgtcaa cataaaatag ctgaaatggt agcacttatg gaaaaacata agcaccaata
2341 tgataagatc attgaagaaa gagactcaga attaggactt tataagagca aagaacaaga
2401 acagtcacat ctgagagcat ctttgagat tgaactatcc aatctcaaag ctgaactttt
2461 gtctgttaag aagcaacttg aaatagaaag agaagagaag gaaaaactca aaagagaggc
2521 aaaagaaaac acagctactc ttaaagaaaa aaaaaaaaaa aaaa

//

se2-1 protein 795aa

1 mekqkpfalf vpprssssqv savkpqtlgg dstffksfnk cteddfefpf aktnlsknge
61 nidsdpalqk vnflpvlegv gnsdchygqg lkdsdlense glsrvyskly keaekikkwk
121 vsteaelrqk esklqenrki ieagrkaige lqfgnekvsl kleegiqenk dlikennatr
181 hlcnlketc arsaektkky eyereetrqv ymdlenniek mitafeelrv qaensrlemh
241 fklkedyeki qhleqeykke indkekqvsl lliqiteken kmkdltfille esrdkvnqle
301 ektklqsenl kqsiekqghhl tkeledikvs lqrsvstgka leedlqiatk ticqlteeke
361 tqmeesnkar aahsfvvtef ettvcsllel lrteqqrlek nedqlkiltm elqkkssele
421 emtkltnnke veleelkkvl geketllyen kqfekiaeel kgteqeligl lqarekevhd
481 leiqltaitt seqyyskev kdlktelenek lknteltshc nklslenkel tqetsdmtle
541 lknqgedinn nkkqeermk qienlqetet qlrneleyvr eelkqkrdev kckldkseen
601 cnnlrkqven knkyieelqg enkalkkkgt aeskqlnvy ikvnklelel esakqkfgei
661 tdtqkeied kkiseenlle evekakviad eavklqkeid krcqhkiaem valmekhkhq
721 ydkieerds elglykskeq eqsslrslasle ielsnlkael lsvkkgleie reekeklkre
781 akentatlke kkkkk

//

Fig. 7 (2)

14/44

se2-2 2317 bp

BASE COUNT : 999 a 334 c 479 g 505 t

ORIGIN :

```
1 ctaccaacaa gcattttatt cgtctggctg agatggaaca gacagtagca gaacaagatg
61 actctctttc ctcaactctt gtcaaaactaa agaaagtatc acaagatttg gagagacaaa
121 gagaaatcac tgaattaaaa gtaaaagaat ttgaaaatat caaattacag cttcaagaaa
181 accatgaaga tgaagtgaag aaagtaaaag cggaagtaga ggattttaaag tatcttctgg
241 accagtcaca aaaggagtca cagtgtttta aatctgaact tcagggtcaa aaagaagcaa
301 attcaagagc tccaacaact acaatgagaa atctagtaga acggctaaag agccaattag
361 ccttgaagga gaaacaacag aaagcactta gtcgggcact tttagaactc cgggcagaaa
421 tgacagcagc tgctgaagaa cgtattatct ctgcaacttc tcaaaaagag gcccatctca
481 atgttcaaca aatcgttgat cgacatacta gagagctaaa gacacaagtt gaagatttaa
541 atgaaaatct tttaaaattg aaagaagcac ttaaaacaag taaaaacaga gaaaactcac
601 taactgataa tttgaatgac ttataataat aactgcaaaa gaaacaaaaa gcctataata
661 aaatacttag agagaaagag gaaattgatc aagagaatga tgaactgaaa aggcaaatta
721 aaagactaac cagtggatta cagggcaaac ccctgacaga taataaacia agtctaattg
781 aagaactcca aaggaaagtt aaaaaactag agaaccaatt agagggaag gtggaggaag
841 tagacctaaa acctatgaaa gaaaagaatg ctaaagaaga attaatagg tggaagaag
901 gtaaaaagtg gcaagccaaa atagaaggaa ttcgaaacaa gttaaaagag aaagagggg
961 aagtctttac ttaacaaag cagttgaata ctttgaagga tctttttgcc aaagccgata
1021 aagagaaact tactttgcag aggaaactaa aaacaactgg catgactgtt gatcaggttt
1081 tgggaatacg agctttggag tcagaaaaag aattggaaga attaaaaaag agaaatcttg
1141 acttagaaaa tgatatattg tatatgaggg ccaccaagc tcttcctcga gattctgttg
1201 tagaagattt acatttacia aatagatacc tccaagaaaa acttcatgct ttagaaaaac
1261 agttttcaaa ggatacatat tctaagcctt caatttcagg aatagagtca gatgatcatt
1321 gtcagagaga acaggagctt cagaaggaaa acttgaagtt gtcactgaa aatattgaac
1381 tgaaatttca gcttgaacaa gcaaataaag atttgccaag attaaagaat caagtcagag
1441 atttgaagga aatgtgtgaa tttcttaaga aagaaaaagc agaagttcag cggaaacttg
1501 gccatgttag agggctctgt agaagtggaa agacaatccc agaactggaa aaaaccattg
1561 gtttaatgaa aaaagtagtt gaaaaagtcc agagagaaaa tgaacagttg aaaaaagcat
1621 caggaatatt gactagtga aaaaatggct atattgagca ggaaaatgaa aaattgaagg
1681 ctgaattaga aaaacttaaa gctcatcttg ggcatcagtt gagcatgcac tatgaatcca
1741 agaccaaagg cacagaaaaa attattgctg aaatgaaag gcttcgtaaa gaacttaaaa
1801 aagaaactga tgctgcagag aaattacgga tagcaaagaa taatttagag atattaaatg
1861 agaagatgac agttcaacta gaagagactg gtaagagatt gcagtttgca gaaagcagag
1921 gtccacagct tgaagggtgt gacagtaaga gctggaaatc cattgtgggtt acaagaatgt
1981 atgaaaccaa gttaaaagaa ttggaaactg atattgcaa aaaaaatcaa agcattactg
2041 accttaacaa gcttgtaaaa gaagcaacag agagagaaca aaaagttaac aaatacaatg
2101 aagaccttga acaacagatt aagattctta aacatgttcc tgaagggtgt gagacagagc
2161 aaggccttaa acgggagctt caagttctta gattagctaa tcatcagctg gataaagaga
2221 aagcagaatt aatccatcag atagaagcta acaaggacca aagtggagct gaaagcacca
```

15/44

2281 tacctgatgc tgatcaacta aaaaaaaaaa aaaaaaa

//

se2-2 protein 761aa

1 meqtvaeqdd slssllvklk kvsqdlqr eitelkvkef eniklqlqen hedevkkvka
61 evedlkylll dsqkesqlk selqaqkean sraptttmrn lverlksqla lkekqqkals
121 rallelraem taaaeeeriis atsqkeahln vqqivdrhtr elktqvedln enllklkeal
181 ktsknrensl tdnlnldlnne lqkkqkaynk ilrekeeidq endelkrqik rltsglqgkp
241 ltdnkqslie elqrkvkkle nqlegkveev dlkpmkekna keelirweeg kkwqakiegi
301 rnklkekege vftltkqlnt lkdlfakadk ekltlqrklk ttgmtvdqvl giraleseke
361 leelkkrrld lendilymra hgalprdsyv edhlqnryl qeklhalekq fskdtyskps
421 isgiesddhc qregelqken lklsseniel kfgleqankd lprlknqvrld lkemceflkk
481 ekaevqrklg hvrgsgrsgk tipelektig lmkkvvekvq reneqlkkas giltsekman
541 ieegeneklka eleklkahlg hqlsmhyesk tkgtekiiae nerlrkelkk etdaaeklri
601 aknnleilne kmtvqleetg krlqfaesrg pqllegadskk wksivvtrmy etklkeletd
661 iakknqsitd lkqlvkeate reqkvnykyne dlegqikilk hvpegaeteq glkrelqvlr
721 lanhqldkek aelihqiean kdqsgaesti pdadqlkkkk k

//

Fig. 8 (2)

se14-3 2620 bp

BASE COUNT: 776 a 733 c 659 g 452 t

ORIGIN

```
1 atcaaatgct gctcgatccc accaacccca gcgcggcac tgccaagata gacaagcagg
61 agaaggtcaa gctcaacttt gacatgacgg catcccccga gatcctgatg agcaagcctg
121 tgctgagtgg gggcacaggc cgccggattt ccttgctcga tatgccgcgc tccccatga
181 gcacaaactc ttctgtgcac acgggctccg acgtggagca ggatgctgag aagaaggcca
241 cgctgagcca cttcagtgcg agcgaggagt ccatggactt cctggataag agcacagctt
301 caccagcctc caccaagacg ggacaagcag ggagtttatc cggcagccca aagcccttct
361 ctctcaact gtcagctcct atcacgacga aaacggacaa aacctccacc accggcagca
421 tcctgaatct taacctggat cgaagcaaag ctgagatgga tttgaaggag ctgagcgagt
481 cgggccagca acagtccacc cctgttcttc tcatctctcc caagcgccag attcgtagca
541 ggttccagct gaatcttgac aagaccatag agagttgcaa agcacaatta ggcataaatg
601 aaatctcgga agatgtctat acggccgtag agcacagcga ttccggaggat tctgagaagt
661 cagatagtag cgatagtgag tatatcagtg atgatgagca gaagtctaag aacgagccag
721 aagacacaga ggacaaagaa ggttgtcaga tggacaaaga gccatctgct gttaaaaaaa
781 agcccaagcc taaaaaccca gtggagatta aagaggagct gaaaagcacg tcaccagcca
841 gcgagaaggc agacctgga gcagtcaagg acaaggccag cctgagcct gagaaggact
901 tttccgaaaa ggcaaacct tcacctcacc ccataaagga taaactgaag ggaaaagatg
961 agacggattc cccaacagtc catttgggcc tggactctga ttcagagagc gaacttgta
1021 tagatthagga agaagaccat tctgggcggg agggtcgaaa aaataagaag gaacccaaag
1081 aaccatctcc caaacaggat gttgtaggta aaactccacc atccacgacg gtgggcagcc
1141 attctcccc ggaaacaccg gtgctcacc gctcttcgc ccaacttcc gcggctggcg
1201 ccacagccac caccagcacg tcctccacgg tcaccgtcac ggccccggcc cccgccgcca
1261 caggaagccc agtgaaaaag cagaggccgc tttaccgaa ggagactgcc ccggccgtgc
1321 agcgggtcgt gtggaactca tcaactgtcc agcagaagga gatcacacag agcccatcca
1381 cgtccaccat caccctggtg accagcacac agtcacggc cctggtcacc agctcggggt
1441 ccatgagcac ccttggtgcc tcagtcaacg ctgacctgcc catcgccact gcctcagctg
1501 atgtcgccgc tgatattgcc aagtacacta gcaaatgat ggatgcaata aaaggaacaa
1561 tgacagaaat atacaacgat ctttctaaaa aactactgg aagcacaata gctgagattc
1621 gcaggctgag gatcgagata gagaagctcc agtggctgca ccagcaagag ctctccgaaa
1681 tgaaacacaa cttagagctg accatggcgg agatgcggca gagcctggag caggagcggg
1741 accggctcat cgccgagggtg aagaagcagc tggagttgga gaagcagcag gcggtggatg
1801 agaccaagaa gaagcagtgg tgcgccaact gcaagaagga ggccatcttt tactgctgtt
1861 ggaacactag ctactgtgac taccctgcc agcaagccca ctggcctgag cacatgaagt
1921 cctgcaccca gtcagctact gtcctcagc aggaagcggg tgctgaggtg aacacagaaa
1981 cactaaataa gtctcccag gggagctcct cgagcacaca atcagcacct tcagaaacgg
2041 ccagcgctc caaagagaag gagacgtcag ctgagaaaag caaggagagt ggctcgacct
2101 ttgaccttcc tggctccaga gagacgacct cctccattct cttaggctcc aaccaaggct
2161 ctgaccattc ccggagtaat aaatccagtt ggagcagcag tgatgagaag aggggatcga
2221 cacgttccga tcacaacacc agtaccagca cgaagagcct cctcccgaaa gagtctcggc
```

2281 tggacacctt ctgggactag cagtgaatcg ggacacaaac caccaccccc attgggagaa
2341 aaacccagac gccaggaaaa gaagaaacaa caaaggcagg agaacagcca ctttcagact
2401 tgaatatgac aaaaccctca gttgagcctg agcccccggc gcgggggctg ctacactaca
2461 ggacacccag catcggttt gactgcagac tgttcaccca cagagccct gtgcttttg
2521 tgtaaataat gtacaatttg tggatgtcat tgaatctaga ggactttccc ctttttatat
2581 ttgtattaac ttttaacttat taaaaaaaaa aaaaaaaaaa

//

sel4-3 protein 764aa

1 mllldptnpsa gtakidkqek vklnfdmtas pkilmskpvl ssgtgrrisl sdmprspms
61 nssvhtgsdv eqdaekkats shfsaseesm dflckstasp astktgqags lsgspkpfsp
121 qlsapittkt dktsttgsil nlnldrskae mdlkelsesv qqqstpvpli spkrqirsrf
181 qlnldkties ckaqlginei sedvytaveh sdsedseksd ssdseyisdd eqskkneped
241 tedkegcqmd kepsavkkkp kptnpveike elkstspase kadpgavkdk aspepekdfs
301 ekakpsphpi kdklkgkdet dsptvhlgl d sdseselvid lgedhsgreg rknkkepkep
361 spkqdvvgkt ppsttvgsht ppetpvltrs saqtsaagat attstsstvt vtapapaatg
421 spvkkqrpll pketapavqr vwnsstvqq keitqspsts titlvtstqs salvtssgsm
481 stlvssvnad lpiatasadv aadiakytsk mmdaikgtmt eiyndlsknt tgstiaeirr
541 lrieieklqw lhqgelsemk hnleltmaem rqsleqerdr liaevkkqle lekqgavdet
601 kkkqwcanc kkaifyccwn tsycdypcqq ahwpehmksc tqsatapqqe adaevntetl
661 nkssqgssss tqsapsetas askeketsae kskesgstld lsgsretpps illgsnqgsd
721 hsrnkssws ssdekrgrstr sdhntststk sllpkersld tfwd

//

Fig. 9 (2)

18/44

se20-4 2830 bp

BASE COUNT 724 a 756 c 814 g 536 t

ORIGIN

```
1 aattcggcac gaggagagct ggttgctga gtctctcag ctctgcttac cggcgcgact
61 agcggcagcg acgcggttaa aagcgaagg gcgagtgcga gtcccctgag ctgtacgaac
121 gcggtcgcca tggaccgcc agatgagggg cctccggcca agaccgcgcg cctgagcagc
181 tccgagtctc cacagcgca cccgcccccg ccgcccgcgc cgcgcgcgct cctccgactg
241 ccgctgcctc caccacagca gcgcccagag ctccaggagg aaacggaggc ggcacaggcg
301 ctggccgata tgaggggggt gggactgggc ccgcgcgtgc cccgcgcgc tccctatgtc
361 attctcgagg agggggggat ccgcgcatac ttacgcgtcg gtgctgagtg tcccggtg
421 gattctacca tcgagtcggg gtatggggag gcgccccgc ccacggagag cctggaagca
481 ctccccactc ctgaggcctc gggggggagc ctggaaatcg attttcagggt tgtacagtcg
541 agcagttttg gtggagaggg ggcctagaa acctgtagcg cagtgggggtg ggcgccccag
601 aggttagttg acccgaagag caaggaagag gcgatcatca tagtgaggga tgaggatgag
661 gatgagcggg agagtatgag gagcagcagg aggcggcggc ggcggcgag gaggaagcag
721 aggaaggtga agagggaaag cagagagaga aatgccgaga ggatggagag catcctgcag
781 gactggagg atattcagct ggatctggag gcagtgaaca tcaaggcagg caaagccttc
841 ctgcgtctca agcgaagtt catccagatg cgaagaccct tcctggagcg cagagacctc
901 atcatccagc atatcccagg cttctgggtc aaagcattcc tcaaccaccc cagaatttca
961 attttgatca accgacgtga tgaagacatt ttccgctact tgaccaatct gcaggtagag
1021 gatctcagac atatctccat gggctacaaa atgaagctgt acttcagac taacccttac
1081 ttcacaaaca tgggtgattgt caaggagttc cagcgcaacc gctcaggccg gctgggtgtct
1141 cactcaacc caatccgctg gcaccggggc caggaacccc aggcgcgtcg tcacgggaac
1201 caggatgcga gccacagctt tttcagctgg ttctcaaacc atagcctccc agaggctgac
1261 aggattgctg agattatcaa gaatgatctg tgggttaacc ctctacgcta ctacctgaga
1321 gaaaggggct ccaggataaa gagaaagaag caagaaatga agaaacgtaa aaccaggggc
1381 agatgtgagg tggatgatcat ggaagacgcc cctgactatt atgcagtgga agacattttc
1441 agcgagatct cagacattga tgagacaatt catgacatca agatctctga cttcatggag
1501 accaccgact acttcgagac cactgacaat gagataactg acatcaatga gaacatctgc
1561 gacagcgaga atcctgacca caatgaggtc cccaacaacg agaccactga taacaacgag
1621 agtgctgatg accacgaac cactgacaac aatgagagtg cagatgacaa caacgagaat
1681 cctgaagaca ataacaagaa cactgatgac aacgaagaga accctaacaa caacgagaac
1741 acttacggca acaacttctt caaagggtggc ttctggggca gccatggcaa caaccaggac
1801 agcagcgaca gtgacaatga agcagatgag gccagtgatg atgaagataa tgatggcaac
1861 gaaggtgaca atgagggcag tgatgatgat ggcaatgaag gtgacaatga aggcagcgat
1921 gatgacgaca gagacattga gtactatgag aaagtatttg aagactttga caaggatcag
1981 gctgactacg aggacgtgat agagatcatc tcagacgaat cagtgggaaga agagggcatt
2041 gaggaaggca tccagcaaga tgaggacatc tatgaggaag gaaactatga ggaggaagga
2101 agtgaagatg tctgggaaga aggggaagat tcggacgact ctgacctaga ggatgtgctt
2161 cagggtccaa acggttgggc caatccgggg aagaggggga aaaccggata agggttttcc
2221 ccttttgggg atcacctctc tgtatcccc acccactatc ccatttgccc tcctcctcag
```

19/44

2281 ctagggccac gcgccccac attgcacttc tggggggtga ccgacttcgt acacggggtt
2341 aaagtttatt tttatggttt agtcattgca gagttcttat tttgggggga gggaaagggg
2401 gctagtcccc ttcttttggc cctccgcccc cgcaggcttc tgtgtgctgc taactgtatt
2461 tattgtgatg ccttggtcag ggccccctcta cccacttctc ccagtcagtt gtggccccag
2521 cccctctccc tgtgctgtgt ggagtggaaca ccctgacccc cgaagcgggg agggccgctg
2581 tggccttcgt cacagccgcg cagtgcccat ggaggcgctg ctgccacctt cctctcccaa
2641 gttctttctc catecctctc ctcttccgc cgcgccgcta gcccgccctg gtgtctatgc
2701 aaggccgctt cgccattgcg gtattctttg cggatttctt gtccccgtcc cccagaaggc
2761 tcgcctctcc ccgtggaccc tgtaaatccc aataaaattc tgagcaagtt caaaaaaaaa
2821 aaaaaaaaaa

//

se20-4 protein 693aa

1 mdrpdegppa ktrrlsses pqrpppppp pppllrplp ppqqrprlqe eteaaqvlad
61 mrgvglgpal pppppyvile eggirayftl gaecpgwdst iesgygeapp pteslealpt
121 peasggslei dfqvvqsssf ggagaletcs avgwapqrlv dpkskeeaii ivedededer
181 esmrssrrrr rrrrrkqrkv kresrernaes rmesilqale diqldleavn ikagkaflrl
241 krkfiqmrrp flerrdliiq hipgfwwkaf lnhprisili nrrdedifry ltnlqvqdlr
301 hismgykmkl yfqtnpyftn mvivkefqrn rsgrlvshst pirwhrgqep qarrhgnqda
361 shsffswfsn hslpeadria eiikndlwnv plryylrerg srikrkkqem kkrktrgrce
421 vvimedapdy yavedifsei sdidetihdi kisdmettd yfettneit dinenicdse
481 npdhnevppn ettdnnesad dhettdnnes addnnped nnkntddnee npnnnentyg
541 nnffkggfwg shgnnqdssd sdneadeasd dedndgnegd negsdddgne gdnegsdddd
601 rdieyyekvi edfdkdqady edvieiisde sveeegieeg iqddediye gnyeeegsed
661 vweegedsdd sdledvlqvp ngwanpgkrk ktg

//

Fig. 10 (2)

20/44

se20-7

2100 bp

BASE COUNT 900 a 329 c 463 g 408 t

ORIGIN

```
1 agagaatcca gaaagtgatg gagagccagt agtgggaagat ggaacttctg taaaaacact
61 ggaaacactc cagcaaagag tgaagcgtca agagaaccta cttgaagcgtt gtaaggaaac
121 aattcagtca cataaggaac aatgtacact attaactagt gaaaaagaag ctctgcaaga
181 acaactggat gaaagacttc aagaactaga aaagataaag gaccttcata tggccgagaa
241 gactaaactt atcactcagt tgcgtgatgc aaagaactta attgaacagc ttgaacaaga
301 taagggaatg gtaatcgag agacaaaacg tcagatgcat gaaaccctgg aaatgaaaga
361 agaagaaatt gctcaactcc gtagtcgcat caaacagatg actaccagg gagaggaatt
421 acgggaacag aaagaaaagt ccgaaagagc tgcttttgag gaacttgaaa aagctttgag
481 tacagcccaa aaaacagagg aagcacggag aaaactgaag gcagaaatgg atgaacaaat
541 aaaaactatc gaaaaaaca gtgaggagga acgcatcagt cttcaacagg aattaagtcg
601 ggtgaaacag gaggttggtg atgtaatgaa aaaatcctca gaagaacaaa ttgctaagct
661 acagaagctt catgaaaagg agctggccag aaaagagcag gaactgacca agaagcttca
721 gacccgagaa agggaatttc aggaacaaat gaaagtagct cttgaaaaga gtcaatcaga
781 atatttgaag atcagccaag aaaaagaaca gcaagaatct ttggccctag aagagttaga
841 gttgcagaaa aaagcaatcc tcacagaaag tgaaaataaa cttcgggacc ttcagcaaga
901 agcagagact tacagaacta gaattcttga attggaaagt tctttgaaa aaagcttaca
961 agaaaacaaa aatcagtcaa aagatttggc tgttcactctg gaagctgaaa aaaataagca
1021 caatatggag attacagtca tggttgaaaa acacaagaca gaattggaaa gccttaagca
1081 tcagcaggat gccctttgga ctgaaaaact ccaagtctta aagcaacaat atcagactga
1141 aatggaaaaa cttagggaaa agtgtgaaca agaaaaagaa acattgttga aagacaaaga
1201 gattatcttc caggccca caagaagaat gaatgaaaag actttagaaa agcttgatgt
1261 gaagcaaaca gaactagaat cattatcttc tgaactgtca gaagtattaa aagcccgta
1321 caaactagaa gaggaacttt ctgttctgaa agatcaaaca gataaaatga agcaggaatt
1381 agaggccaag atggatgaac agaaaaatca tcaccagcag caagttgaca gtatcattaa
1441 agaacacgag gtatctatcc agaggactga gaaggcatta aaagatcaaa ttaatcaact
1501 tgagcttctc ttgaaggaaa gggacaagca tttgaaagag catcaggctc atgtagaaaa
1561 tttagaggca gatattaaaa ggtctgaagg ggaactccag caggcatctg ctaagctgga
1621 cgtttttcag tcttaccaga gtgccacaca tgagcagaca aaagcatatg aggaacagtt
1681 ggcccaattg cagcagaagt tgttggaatt ggaaacagaa agaattcttc ttaccaaaaca
1741 ggttgctgaa gttgaagcac aaaagaaaga tgtttgact gagttagatg ctcacaaaat
1801 ccaggtgcag gacttaatgc agcaacttga aaaacaaaat agtgaaatgg agcaaaaagt
1861 aaaatcttta acccaagtct atgagtccaa acttgaagat ggtaacaaag aacaggaaca
1921 gacaaagcaa atcttggtg aaaaggaaaa tatgatttta caaatgagag aaggacagaa
1981 gaaagaaatt gagatactca cacagaaatt gtcagccaag gaggacagta ttcatathtt
2041 gaatgaggaa tatgaaacca aatttaaaaa aaaaaaaaaa aaaaaaaaaa aaaaaaaaaa
```

//

Fig. 11 (1)

21/44

se20-7 protein

623aa

1 maektklitq lrdaknlieq leqdkgmvia etkrqmhetl emkeeeiaql rsrikqmttq
61 geelreqkek seraafeele kalstaqkte earrklkaem deqiktiekt seeerislqq
121 elsrvkqevv dvmkksseeq iaklqklhek elarkeqelt kklqtreref qeqmkvalek
181 sqseylkisq ekeqqeslal eelelqkkai ltesenklrd lqqaetyrt rilelessle
241 kslqgenknqs kdlavhleae knkhnmeitv mvekhktele slkhqqdalw tekqlvklqq
301 yqtemekltre kceqeketll kdkeiifqah ieemnektle kldvkqtele slsselsevl
361 karhkleeel svlkdqtdkm kqeleakmde qknhhqqqvvd siikehevsi qrtekalkdq
421 inqlellllke rdkhlkehqa hvenleadik rsegelqqas akldvfqsyq satheqtkay
481 eeqlaqlqqk lldleteril ltkqvaevea qkkdvcteld ahkiqvqdlm qqlekqnsem
541 eqkvksltqv yeskledgnk egeqtkqilv ekenmilqnr egqkkeieil tqklsakeds
601 ihilneeyet kfkkskkkkk kkk

//

Fig. 11 (2)

se20-9 5629 bp

BASE COUNT 2030 a 905 c 1181 g 1513 t

ORIGIN

```
1 ctttaagtgc aggaactctg tgttgggagg aaaaatgtcc ttcttcaatt tccgtaagat
61 cttcaagttg gggagcgaga agaagaagaa gcagtagcaa cacgtgaaga gggacctgaa
121 ccccgaagac ttttgggaga ttataggaga actgggagac ggagcctttg ggaaagtgtg
181 caaggcccag aataaagaga ccagtgtttt agctgctgca aaagtgattg acactaaatc
241 tgaagaagaa cttgaagatt acatggtaga gattgacata ttagcatctt gtgatcacc
301 aatatagtc aagcttctag atgccttcta ttatgagaac aatctttgga tcctcattga
361 attttgtgca ggtggagcag tagatgctgt gatgcttgaa cttgagagac cattaactga
421 gtcccaaata caagtagttt gcaagcagac ttagatgca ttgaactact tacatgataa
481 taagatcatc cacagagatc tgaaggctgg caacattctc ttacottag atggagatat
541 caaattggcg gattttggag tatcagctaa aaacacgagg acaattcaaa gaagagattc
601 ctttattggt acaccatatt ggatggctcc tgaagtagtc atgtgtgaaa catctaagga
661 cagaccctat gactacaaag ctgatgtttg gtccctgggt atcactttaa tagaaatggc
721 tgagatagaa ccacctcatc atgaattaaa tccaatgcga gtgctgctaa aaatagcaaa
781 atctgagcca cctacattag cacagccatc cagatggctt tcaaatttta aggactttct
841 aaagaaatgc ttagaaaaga atgtggatgc caggtggact acatctcagc tgctgcagca
901 tccctttggt actgttgatt ccaacaaacc catccgagaa ttgattgcag aggcgaaggc
961 tgaagtaaca gaagaagttg aagatggcaa agaggaagat gaagaggagg aaacagaaaa
1021 ttctctgcca atacctgcaa gtaagcgtgc atcttctgac cttagtatcg ccagctctga
1081 agaagataaa ctttcacaaa atgcttgat tttggagtct gtctcagaaa aaacagaacg
1141 tagtaactct gaagataaac tcaacagcaa aattcttaat gaaaaacca cactgatga
1201 acctgaaaag gctgtggagg atattaatga acatattacc gatgctcagt tagaagcaat
1261 gactgaactc catgacagaa cagcagtaat caaggagaat gaaagagaga agaggcccaa
1321 gcttgaaaat ctgcctgaca cagaagacca agaaactgtg gacattaatt cagtcatgga
1381 aggaaaagag aataatataa tgataacctt agaaacaaat attgaacata atctaaaatc
1441 tgaggaagaa aaggatcagg aaaagcaaca gatgtttgaa aataagctta taaaatctga
1501 agaaattaaa gatactatct tgcaaacagt agatttagtt tctcaagaga ctggagaaaa
1561 agaggcaaat attcaggcag ttgatagtga agttgggctt acaaaggaag acaccaaga
1621 gaaattgggg gaagacgaca aaactcaaaa agatgtgatc agcaatacaa gtgatgtgat
1681 aggaacatgt gaggcagcag atgtggctca gaaagtggat gaagacagtg ctgaggatac
1741 gcagagtaat gatgggaaag aagtggctga agtaggccag aaattaatta ataagcccat
1801 ggtgggtcct gaggtgtgtg gtactaagga agttcctatt aaagaaatag ttgaaatgaa
1861 tgaaatagaa gaaggtaaaa ataaggaaca agcaataaac agttcagaga acataatgga
1921 catcaatgag gaaccaggaa caactgaagg tgaagaaatc actgagtcaa gtagcactga
1981 agaaatggag gtcagaagtg tgggtggctga tactgaccaa aaggctttag gaagtgaagt
2041 tcaggatgct tctaaagtca ctactcagat agataaagag aaaaaagaaa ttccagtgtc
2101 aattaaaaaa gagcctgaag ttactgtagt ttcacagccc actgaacctc agcctgttct
2161 aatacccagt attaatatca actctgacag tggagaaaat aaagaagaaa taggttcttt
2221 atcaaaaact gaaactattc tgccaccaga atctgagaat ccaaaggaaa atgataatga
```

23/44

2281 ttcaggcact ggttcactg ctgatactag cagtattgac ttgaatttat ccatctctag.
2341 ctttctaagt aaaactaaag acagtggtatc gatattcttta caagaaacaa gaagacaaaa
2401 gaaaacattg aagaaaacac gcaaatttat tgttgatggt gtagaagtga gtgtaacaac
2461 atcaaagata gttacagata gtgattccaa aactgaagaa ttgcggtttc ttagacgtca
2521 ggaacttcgg gaattaagat ttcttcagaa agaagagcaa agagcccaac aacagctcaa
2581 tagcaaaacta cagcaacaac gagaacaaat tttccggcgc tttgagcagg aaatgatgag
2641 taaaaagcga caatatgacc aggaaattga gaatctagaa aaacagcaga aacagactat
2701 cgaacgcctg gaacaagagc acacaaatcg cttgcgagat gaagccaaac gcatcaaagg
2761 agaacaagag aaagagttgt ccaaatttca gaatatgctg aagaaccgaa agaaggaggt
2821 tataaatgaa gtggagaaag cacccaaaga gctgagaaaa gagctcatga aacgcaggaa
2881 agaggagctt gcacaaagcc agcatgctca ggaacaagag tttgttcaga aacaacagca
2941 agaattagat ggctctctga aaaagatcat ccagcagcag aaggcagagt tagctaatat
3001 tgagagagag tgcctgaata acaagcaaca gctcatgaga gctcgagaag ctgcaatttg
3061 ggagctcgaa gaacgacact tacaagaaaa acaccagctg ctcaaacagc agcttaaaga
3121 tcagtatttc atgcaaagac atcagctact taagcgccac gagaaggaaa cagagcaaat
3181 gcagcgttac aatcaaagac ttattgagga attgaaaaac agacagactc aagaaagagc
3241 aagactgccc aagattcagc gcagtgaagc caagactcga atggccatgt ttaagaagag
3301 tttgagaatt aactcaacag ccacaccaga tcaggaccgt gataaaaatta aacagtttgc
3361 tgcacaagaa gaaaagaggc agaaaaatga gagaatggct cagcatcaga aacatgagaa
3421 tcaaattgca gatcttcagt tgcagtgtga agccaatgtc cgcgaactgc atcagctgca
3481 gaatgaaaaa tgccacttgt tggttgagca tgagactcag aaactgaagg agttagatga
3541 ggaacatagc caagaattaa aggagtggag agagaaattg agacctagga aaaagacact
3601 ggaagaagag tttgccagga aactacagga acaggaagta ttctttaaaa tgactgggga
3661 gtctgaatgc cttaacccat caacacagag ccggttttcc aaattttatc ctattcccag
3721 cttgcattcc accggatcat aacaaagga agcattctgt gcgtgggttt ggctctttca
3781 gtatgtcatt ctgttctcat cttctgccac agtctctcag atagctcatg aagacaatca
3841 cctgcctcac cttctaggtg ttttctttt ttgtttttt ttgtttgttt tgtttttaag
3901 caaagatgaa gggaaaacga actaagacag acgctaggcc atgttggaag agtagcatct
3961 tggtgactaa ggtgactttg tatattcatc ttaaaaaatta tgttcttttag acactgctac
4021 ctgaaaactg ttggagaaat aatgttttaa gttatttaag aaaaactggt acatcactaa
4081 gtattaataa attcttctta cctgacgtaa cttctcaatg cctaaattct gtagttgaag
4141 ctctgctgca gagagtggg ataattttct tttgggtgat cagctctcat aaaaaagcta
4201 tgatttgctc aaatatgctg ttgactcagt aaatgaatat atttttttct ttaaatagga
4261 acaacctctt ttaaaagaga aaaattattt cagtgatttg tcaaaacgaa ttacctctt
4321 tggcatgagc taataattga gggtgctaatt tttcttaaga tagtgccata aacactaaat
4381 ttcagtcaag tcgtaagtag gattttcttt ttgatcaaca gggacaaaaa catctttaga
4441 attaaaaaca tggttgtttt ggaatttttg cttctcttac cgtttgatag aaattttcat
4501 cctaaaatac atgtacaaag tttggaaaga tgaaaaaag aggtagcttt tagattgcaa
4561 attggaaatg taaaactcat gaaatttaag caatataggt ttagctatct gtgtttattt
4621 tctaaaataa tacctgagct ggttaaatga tttctctcca tcttagctaa ttctgtttaa
4681 aactctgtca gaggcctgca ggctgtgagt tatatttata aatatatctt cagaaattaa

24/44

4741 tcttaaaaga ggcattagtt cagaatactt ttttaaaagt ttaaattaaa tatttaggca
 4801 cgtcagaaat tacttttctt tattttgaaa tgaggctact tatgtcttgg ttttattttg
 4861 ttccatgttt aaatcattca ctttgatttg agtgggaaaa gcctgaagcc tttatcatgt
 4921 ggttgctggg gtgtgtaatt attaatgaaa tgttcaactc tagtccctta tgaggcttag
 4981 aatttcaacc acgtgtcagg tcagacagta ttataaactg tactttgctg tctgagacag
 5041 cacatttgtg aatgatgctt gctgcctgcc attttcaacc tattctctct taagagtgt
 5101 aggtacccaa ttgtgaaagt ttgttttcag ttatattact tttgaggctg gtgaaaaatt
 5161 taaatgtaac tttgtgggaa cactgattca tatttagaaa atgtaaattg ctgtagcact
 5221 ttcttgcaat taatttgaaa actttggatg ctgaaccttg tttgtcagtg atttagatga
 5281 tttaaaaatg catgtgtgat ttgaatttta taattgtttt gacaagcata atttacttgg
 5341 acaacttcgt aggtagcctt aacttctggc caagtttgtt tttatataaa atatataac
 5401 atatatacat attatgtatg gttgtaaatt catacactta tcacatgaat gtgttactgt
 5461 atacaaaact cttaatgctt tattctcaaa tgctgggttg aaaaatgttt tgaaagcctt
 5521 ttaaaatata tatctttata aagtaatatt caggatgatg ataaaaattg tttatattgt
 5581 tatgataaaa atgacagtat aatgttaaaa aaaaaaaaaa aaaaaaaaaa

e20-9 protein

1235aa

1 msffnfrkif klgsekkkkq yehvkrdlp edfweiligel gdgafgkvyk aqnketsvla
 61 aakvidtkse eeledymvei dilascdhpn ivklldafyy ennlwilief caggavdavn
 121 lelerpltes qiqvvckqtl dalnylhdkn ihhrldkagn ilftldgdik ladfgvsakn
 181 trtiqrrdsf igtpywmapv vvmcetskdr pydykadvws lgitliemae iepphheln
 241 mrvllkiaks epptlaqpsr wssnfkdflk kcleknvdar wttsqllqhp fvtvdsnkpi
 301 reliaekae vteeedgke edeeetens lpipaskras sdlsiassee dklsqnacil
 361 esvsekters nsedklnski lnekpttdep ekavedineh itdaqleamt elhdrtavik
 421 enerekprkl enlpdtedge tvdinsvseg kennimitle tniehnlkse eekdqekqgm
 481 fenkliksee ikdtilqtd lvsqetgeke aniqavdsev gltkedtqek lgeddktqkd
 541 visntsdvig tceaadvaqk vdedsaedtq sndgkevvev gqklinkpmv gpeaggtkev
 601 pikeivemne ieegknkeqa inssenimdi neepgttege eitesstee mevrsvvad
 661 dqkalgsevg daskvttqid kekkeipvsi kkepevtvvs qptepqpqli psininsdsg
 721 enkeeigsls ktetilppes enpkendnds gtgstadtss idlnlsissf lsktkdsgsi
 781 slqetrrqkk tlkkrkfiv dgvevsvtts kivtdsdskt eelrflrrqe lrelrflqke
 841 egragqqqlns klqqqreqif rrfegemmsk krqydgeien lekqqkqtie rlegehtnrl
 901 rdeakrikge gekelskfqn mlknrkkevi nevekappel rkelmkrrke elaqsqhage
 961 qefvqkqqqe ldgslkkiiq qqkaelanie reclnnkqql mrareaaawe leerhlqekh
 1021 qllkqqlkdq yfmqrhqlk rheketeqmq rynqrliel knrqtqerar lpkiqrseak
 1081 trmamfkksl rinstatpdq drdkikqfaa geekrqkner maqhkheng mrdlqlqcea
 1141 nvrelhqlqn ekchllvehe tqklkeldee hsqelkewre klrprkktle eefarklqeg
 1201 evffkmtges eclnpstqsr iskfypipsi hstgs

Fig. 12 (3)

se33-1. 3962 bp

BASE COUNT 1455 a 627 c 867 g 1013 t

ORIGIN

```
1 aattcggcac gaggatgagt atagggtgtg ttcttgcagg gttctttcga aattattgct
61 aggttgactt ttaactaaat ccaagtgatg ttatttgtaa tgtagtactt aaatgttttt
121 cttgttggtt tagccaaaac tggacaagcc aaggcatctg tagccaaagt aaacaaatct
181 acagggaaat cagcaagttc tgtaaaatct gtggtaacgg tagctgttaa aggtaataaa
241 gcttcaatca aaacagcaaa atctggtgga aagaagtctc tagaagccaa aaagactggg
301 aatgtcaaaa acaaagactc taacaaacct gtgactatac cagaaaactc tgaaataaag
361 accagtattg aagtcaaagc cactgaaaac tgtgctaaag aagctatttc tgatgctgct
421 ttggaggcca cagagaatga accacttaac aaggaaacag aagaaatgtg tgtgatgctt
481 gtctctaatt tgcctaataa aggatattct gtagaagaag tttatgactt agcaaaacca
541 tttggtggtt taaaggatat cttgatttta tcatctcata aaaaggcata tatagaaata
601 aatagaaaag ctgctgagtc tatggtaaaa ttttatacct gcttcccagt attgatggat
661 ggaaatcaac tctcaataag tatggctcct gaaagcatga atataaaaga tgaggaagct
721 atattttataa ccttggtaaa agaaaatgac ccagaggcaa acatagatac aatttatgat
781 cgatttgtag atcttgataa ttaccggaa gatggacttc agtgtgtact ttgtgttgga
841 cttcagtttg gaaaagtgga tcaccatgta ttcataagta atagaaacaa ggcaattctt
901 cagttagata gtcctgaatc tgctcagtc atgtatagct ttctgaaaca aaatccacaa
961 aatattggtg accatatgtt gacctgctca ttatctccaa agatagactt accagaggtg
1021 caaattgagc atgaccacga attagaaaaa gaaagccctg gcttgaaaaa cagtccaatt
1081 gatgaaagtg aggtgcaaac agcaactgat agtccctctg ttaaacctaa tgagcttgaa
1141 gaagaaagta ctcccagcat tcaaacagaa actttggtac agcaggaaga gccttgtag
1201 gaagaagctg aaaaagcaac atgtgattct gactttgctg ttgaaacttt ggagcttgaa
1261 actcaaggag aggaggtcaa agaagaaatt cctcttgtag catccgcttc agtcagtatt
1321 gaacaattca ctgaaaatgc cgaggagtgt gctttaaatc agcagatgtt taacagtgc
1381 ttggagaaga aaggggcaga aattattaac cctaaaacag cattgttacc atctgacagt
1441 gtgtttgcag aagaaaggaa cctcaaagga attctagaag aatctccatc tgaagcagaa
1501 gatttcattt ctggaattac acagactatg gtagaagctg tagctgaagt agaaaaaaat
1561 gaaactgttt cggaatatt gccatcaact tgtattgtga cgtagtacc aggaattccc
1621 actgggggatg agaagacagt ggacaaaaag aatatttctg aaaaaaagg taacatggat
1681 gaaaaggagg agaaggaatt taatactaag gaaaccagaa tggatcttca aataggaaca
1741 gagaaggctg aaaagaatga aggtaggatg gatgcagaaa aggtggaaaa gatggcagca
1801 atgaaagaaa agcctgcaga aaacacttta ttcaaggcat acccaaataa aggagtgggt
1861 caggctaata agcctgatga aactagtaaa actagtattc tggctgtatc agatgtatct
1921 agcagtaaac caagcatcaa ggctgttata gtctcttctc ctaaggcaaa agctacagtt
1981 tcaaaaactg aaaatcagaa agtttttcca aaatctgtgc ccagagatca aataaatgct
2041 gaaaagaaac tttcagccaa ggaatttggg ctgcttaaac ccacaagtgc caggtcaggc
2101 ttggcagaaa gcagcagtaa attcaaacct actcagagca gtcttaccag aggaggcagt
2161 ggaaggatct cagccctgca aggcaagctt tctaaactgg attacagaga tataacaaaa
2221 caatctcagg aaacagaggc tagaccttcc atcatgaaac gggatgacag caacaataag
```

2281 acttttgctg agcaaacac taagaatcct aaaagcacta ctggtagaag ttccaaatct
2341 aaagaggagc cattatttcc atttaatttg gatgaatttg ttactgtgga tgagggtata
2401 gaagaagtga atcctttctca ggccaagcag aatccactaa agggaaaaag gaaagaaact
2461 ctcaaaaatg ttcctttctc tgaacttaac ttaaagaaga aaaaggggaa aacttccact
2521 cctcgtggtg ttgagggaga actatctttt gtgacattgg atgagattgg ggaagaggaa
2581 gatgcagctg cacatctagc acaagctcta gtcactgtgg atgaagtaat tgatgaagaa
2641 gaactaaata tggaagaaat ggtaaaaaat tcaaattcac tttttacatt agatgaatta
2701 attgaccaag atgattgcat ttcccacagt gaacctaaag atgttactgt tctgtcagtg
2761 gctgaagaac aagatctcct caaacaggaa cgcttggtta ctgtggatga aattggagaa
2821 gtggaagagc tacctttgaa tgagtcagca gacataactt ttgccacttt aaatactaaa
2881 ggaaatgaag gagatatcgt aagggtattcc attggcttca tttcttctca ggtgcccgaa
2941 gacccttcta ctttagttac tgtagatgaa atacaagatg acagcagtga tttgcattta
3001 gtgacttttg atgaagtaac tgaagaggat gaagactctc tggcggattt taacaacctt
3061 aaagaagagc ttaattttgt tactgttgat gaagttggag aggaggaaga tggagataat
3121 gatttaaaaag ttgagttagc acaaagcaaa aatgaccatc ccacagataa aaaagggaat
3181 agaaagaaga gagctgtgga cacaaaaaag acaaaacttg aatccttgtc ccaagtgggt
3241 ccagtaaatg agaattgtat ggaagaagat ctaaaaacca tgattgaaag acacttaaca
3301 gctaaaaactc caaccaagag agttagaatt gggaaaactc tgccatcaga aaaagctggt
3361 gtgacagaac cagcaaaagg tgaagaggcc ttccagatga gtgaagtga tgaggaatct
3421 ggattaaagg attcagaacc agagcgaaaa cgcaagaaga ctgaagactc ttcttcaggc
3481 aaatcagtgg tgtctgatgt ccctgaggaa ttagactttc ttgtacctaa ggctggatc
3541 ttctgtccaa tttgttcctt cttctactca ggtgaaaaag caatgacaaa tcaactgcaag
3601 agtacacgtc ataagcaaaa tactgagaaa ttcatggcca agcaaagaaa ggaaaaggag
3661 cagaatgagg ctgaagaaag aagctctagg tgattggggg aaaggaaaga attcactaga
3721 aatttgttta ggggtccagt gatttgtgta tttttgttat catttaattt gtaattttcg
3781 tttcagaagc aaatattcgt gttgtacaaa tttctgattg ccctaaatgt agagagactg
3841 atggggaaag tatgatgggt ttgattttta tatcaaatca tcaggcatgg agaaatatct
3901 tttagaagtg ttaaaataaa tgttcctact gtatatttaa aataaaaaaa aaaaaaaaaa
3961 aa

//

Fig. 13 (2)

se33-1 protein 1075aa

27/44

1 mcvmlvsnlp nkgysveevy dlakpfggk dililsshkk ayieinrkaa esmvkfytcf
61 pvlmdgnqls ismapesmni kdeeaifitl vkendpeani dtiydrfvhl dnlpedglqc
121 vlcvglqfgk vdhhvfisnr nkailqldsp esaqsmysfl kqnpqnigdh mlcslspki
181 dlpevqiehd pelekespgl knspidesev qtatdspsvk pneleestp siqtetlvqq
241 eepceeeaaek atcdsdfave tleletqgee vkeeiplvas asvsieqfte naeecalnqq
301 mfnsllekkg aeiinpkta lpsdsvfaee rnlkgilees pseaedfisg itqtmveava
361 eveknetvse ilpstciavl vpgiptgdek tvdkknisek kgmdekeek efntketrmd
421 lqigtekaek negrmdaekv ekmaamkekp aentlfkayp nkgvggankp detsktsila
481 vsdvssskps ikavivsspk akatvskten qksfpksvpr dqinaekkl akefgllkpt
541 sarsglaess skfkptqssl trggsgrisa lqgklskldy rditkqsqet earpsimkrd
601 dsnkntlaeq ntknpksttg rsskskeepl ffnldefvt vdevieevnp sqakgnplkg
661 krketlknvp fselnlkkkk gktstprgve gelsfvtlde igeedaaaah laqalvtvde
721 videeelme emvknsnsf tldelidqdd cishsepkdv tvlsvaeeqd llkqerlvtv
781 deigeveelp lnesaditfa tlntkgnegd ivrdsigfis sqvpdpstl vtvdei qdds
841 sdlhlvtlde vteededs la dfnnlkeeln fvtvdevgee edgdndlkve laqskndhpt
901 dkkgnrkkra vdtkktkles lsqvgpvnen vmeedlktmi erhltaktpt krvrigktlp
961 sekavvtepa kgeaafqmse vdeesglkds eperkrkkte dsssgksvvs dvpeeldflv
1021 pkagffcpic slfysgekam tnhckstrhk qntekfmakq rkekeqneae erssr

//

Fig. 13 (3)

se37-2 2710 bp

BASE COUNT 930 a 468 c 584 g 728 t

ORIGIN

```
1 ggagatacaa gtttggaagc aatcttgggg tacttaccca caaggctggt ggagaccaga
61 tcaggagaac ctcaagtctga cgacattgaa gctagccgaa tgaagcgagc agctgcaaag
121 catctaataag aacgctacta ccaccagtta actgagggct gtggaaatga agcctgcacg
181 aatgagtttt gtgcttcctg tccaactttt cttcgtatgg ataataatgc agcagctatt
241 aaagccctcg agctttataa gattaatgca aaactctgtg atcctcatcc ctccaagaaa
301 ggagcaagct cagcttacct tgagaactcg aaaggtgccc ccaacaactc ctgctctgag
361 ataaaaatga acaagaaagg cgctagaatt gatttttaaag atgtgactta cttaacagaa
421 gagaaggat atgaaattct tgaattatgt agagaaagag aggattattc ccctttaatc
481 cgtgttattg gaagagtttt ttctagtgtg gaggcattgg tacagagctt ccggaagtt
541 aaacaacaca ccaaggaaga actgaaatct cttcaagcaa aagatgaaga caaagatgaa
601 gatgaaaagg aaaaagctgc atgttctgct gctgctatgg aagaagactc agaagcatct
661 tcctcaagga taggtgatag ctcacaggga gacaacaatt tgcaaaaatt aggccctgat
721 gatgtgtctg tggatattga tgccattaga aggtctaca ccagattgct ctctaataaa
781 aaaattgaaa ctgcctttct caatgcactt gtatatttgt cacctaacgt ggaatgtgac
841 ttgacgtatc acaatgtata ctctcgagat cctaattatc tgaatttggt cattatcgta
901 atggagaata gaaatctcca cagtcctgaa tatctggaaa tggctttgcc attattttgc
961 aaagcgatga gcaagctacc ccttgagccc caaggaaaac tgatcagact gtggtctaaa
1021 tacaatgcag accagattcg gagaatgatg gagacatttc agcaacttat tacttataaa
1081 gtcataagca atgaatttaa cagtcgaaat ctagtgaatg atgatgatgc cattgttgct
1141 gcttcgaagt gcttgaaaat ggtttactat gcaaatgtag tgggagggga agtggacaca
1201 aatcacaatg aagaagatga tgaagagccc atccctgagt ccagcgagct gacacttcag
1261 gaacttttgg gagaagaaag aagaaacaag aaaggtcctc gagtggaccc cctggaaact
1321 gaacttggtg ttaaaacctt ggattgtcga aaaccactta tcccttttga agagtattt
1381 aatgaaccac tgaatgaggt tctagaaatg gataaagatt atactttttt caaagtagaa
1441 acagagaaca aattctcttt tatgacatgt ccctttatat tgaatgctgt cacaaagaat
1501 ttgggattat attatgacaa tagaattcgc atgtacagtg aacgaagaat cactgttctc
1561 tacagcttag ttcaaggaca gcagttgaat ccatatttga gactcaaagt tagacgtgac
1621 catatcatag atgatgcact tgtccggcta gagatgatcg ctatggaaaa tcctgcagac
1681 ttgaagaagc agttgtatgt ggaatttgaa ggagaacaag gagttgatga gggaggtgtt
1741 tcaaagaat tttttcagct ggttgtggag gaaatcttca atccagatat tggatgttc
1801 acatacatg aatctacaaa attgttttgg ttaaatccat cttcttttga aactgaggtt
1861 cagtttactc tgattggcat agtactgggt ctggctatct acaataactg tatactggat
1921 gtacattttc ccatggttgt ctacaggaag ctaatgggga aaaaagggaac ttttcgtgac
1981 ttgggagact ctaccccagt tctatatcag agtttaaaag atttattgga gtatgaaggg
2041 aatgtggaag atgacatgat gatcacttcc cagatatcac agacagatct ttttggtaac
2101 ccaatgatgt atgatctaaa ggaaaatggt gataaaaatc caattacaaa tgaaaacagg
2161 aaggaatttg tcaatcttta ttctgactac attctcaata aatcagtaga aaaacagttc
2221 aaggcttttc ggagagggtt tcatatggtg accaatgaat ctcccttaaa gtacttattc
```


2281 agaccagaag aaattgaatt gcttatatgt ggaagccgga atctagattt ccaagcacta
2341 gaagaaacta cagaatatga cggtaggctat accagggact ctgttctgat tagggagttc
2401 tgggaaatcg ttcattcatt tacagatgaa cagaaaagac tcttcttgca gtttacaacg
2461 ggccacagaca gagcacctgt gggaggacta ggaaaattaa agatgattat agccaaaaat
2521 ggcccagaca cagaaagggtt acctacatct catacttgct ttaatgtgct ttacttccg
2581 gaatactcaa gcaaagaaaa acttaaagag agattgttga aggccatcac gtatgccaaa
2641 ggatttgga tgctgtaaaa caaaacaaaa caaataaaaa caaaaaaaaa aaaaaaaaaa
2701 aaaaaaaaaa

//

se37-1 protein

852aa

1 mkraaakhli eryyhlteg cgneactnef cascptflrm dnnaaaikal elykinaklc
61 dphpskkgas saylenskga pnnscseikm nkkgaridfk dvtylteekv yeilelcrer
121 edysplirvi grvfssaeal vqsfkrvkqh tkeelkslqa kdedkdedek ekaacsaaam
181 eedseasssr igdssqgdnn lqklgpddvs vdidairrvy trllsnekie taflnalvyl
241 spnvecdlty hnvysrdpny lnlfiiivmen rnlhspeyle malplfckam sklplaaqgk
301 lirlwskyna dqirrmmtf qqlitykvis nefnsrnlvn ddaivaask clkmvyyanv
361 vggevdtahn eeddeepipe sseltlqell geerrnkkgp rvdpletelg vktldcrkpl
421 ipfeefinep lnevlemdkd ytffkveten kfsfmtcpfi lnavtknlgl yydnrirmys
481 erritvlysl vggqqlnpyl rlkvrrdhii ddalvrlemi amenpadlkk qlyvefegeq
541 gvdeggvske ffqlvveeif npdigmftyd estklfwfnp ssfetegqft ligivlgiai
601 ynnchildvhf pmvvyrkmg kkgtfrdlgd shpvlyqslk dlleyegnve ddmmitfqis
661 qtdlfgnppm ydlkengdki pitnenrkef vnlysyiln ksvekqfkaf rrgfhmvtne
721 splkylfrpe eiellcgsr nldfqaleet teydggytrd sylirefwei vhsftdeqkr
781 lflqfttgd rapvgglgkl kmiiakngpd terlptshtc fnvlllpeys skeklkerll
841 kaityakgfg ml

//

Fig. 14 (2)

This Page Blank (uspto)

se89-1. 3462 bp

BASE COUNT 1195 a 580 c 688 g 999 t

ORIGIN

```
1 cagatgtatt aaaaatagct cagtttttac aattttcttt gattcagtgt cgaaaggaat
61 tcaaaaatat aagcgccata aattttcatt ctgttggtga aaagtatgta agtgaatttt
121 ttaagcgagg ttttggttca ggtaaacgag agtttattat gtttccatat gattcacgat
181 tagatgataa aaaattctta tactcagctc ccagaaataa atcccatatt gatacttggt
241 tgcattgcta tatttttcgg cctgaagtgt atcagttacc tatttgtaaa ttaaaagaac
301 tatttggaaga aaatagaaaa cttcagcagt ttagtccact ttcagattat gaaggccaag
361 aagaagaaat gaatggtaca aaaatgaaat ttggaaaacg aaataactca agagggtgaag
421 ccattatata tggaaagcaa agatcatctc attctttgga ttatgataag gatagagtca
481 aagaattgat taatttaatt cagtgtagga aaaagagtgt ggggtggggac tcagacacag
541 aagatatgag aagcaaaact gtcttgaaga ggaagcttga ggatctacct gaaaatatga
601 gaaagctcgc caaaaccagt aatttatctg aaaattgcca tctgtatgaa gagtctccac
661 agcctattgg ctcaattgga catgatgctg acttgaggcg gcagcagcag gatacctgta
721 actccggcat tgctgacatc cataggctgt ttaattgggt atcagaaaca ctagcaaatg
781 cgcgccattc tgatgcatct ctgacagaca cagtcaacaa agccttagga ttgagcactg
841 atgatgcta tgaagagctg aggcaaaaac atgagtatga gttgaactct accccagata
901 agaaagacta tgagcagcct acttggtgcaa aagttgaaaa tgcacagttt aagggtactc
961 agagcttatt actagaagtt gatgcaacat ctaagtattc tgttgctatt tctaccagcg
1021 aagtgggcac tgaccataag ctacatttga aagaagatcc aaatttaatt agcgtgaata
1081 attttgaaga ttgcagtttg tgtcccagtg ttcccattga acatggattt cgtagacaac
1141 agtctaagtc aaataatggt gaagagactg aaatacattg gaaactgatt ccaattacag
1201 acacactaaa gggcaccact gaggatgacg tgttgacagg tcagggtggag gagcagtggtg
1261 tgccagcagc agaggcagag ccgcctgcag tgagcgaaac cacagagagg acagtgttag
1321 gagagtacaa tctcttttct aggaagatag aagagatttt gaagcaaaaag aatgtttcat
1381 atgtcagtag agtttccaca cctatctttt caacacaaga gaagatgaaa cggctttccg
1441 agttcatata ttctaagact tccaaagctg gtgtgcagga gttttagatg ggtttgcatg
1501 agaagctaaa tactattatt attaaagcat cagccaaggg tgggaatttg ccaccagtca
1561 gtcctaacga ttctggtgct aagatagcat cgaatcctct ggaaaggcat gtcataccag
1621 tttcctcaag tgacttcaac aataaacatc tccttgagcc actgtgtagt gatcctttga
1681 aagataccaa ctctgatgag cagcattcca cttcggtttt aactgaagta gaaatgaacc
1741 agcctcaaca tgccacagag ttaatggtga cttctgatca tattgtacct ggtgatatgg
1801 cccgggaacc agtagaagaa acaacaaaat ccccagtgta tgtaaacatt tctgctcagc
1861 cagctctttc aaattttata agccagttag aacctgaagt atttaatagt ttggttaaaa
1921 tcatgaaaga cgtccagaaa aatactgtga aattttatat tcatgaagaa gaagagagtg
1981 tgctctgtaa agaaataaag gaatatctta tcaaattagg caatacagaa tgtcatcctg
2041 aacagttttt ggaaagaaga tcaaaattag ataaactatt gattattatt caaaatgaag
2101 acattgcagg tttcaattcac aagatacctg gcttggtgac tttaaagaag ctcccctgtg
2161 ttagttttgc tgggtgtgat agcctggatg atgttaaaaa tcatacatc aatgaattat
2221 ttgtatctgg aggttttata gtatctgatg aatcaattct aaaccagag gttgtcacag
```

2281 ttgagaacct taaaaatttt ttgacattcc ttgaggaact tagtactcca gaaggaaaat
2341 ggcaatggaa agtccactgt aaatttcaga agaaactaaa ggaactaggg agattgaatg
2401 cttaaagctct aagtctgttg acgcttctga atgtctatca gaagaaacat ctggttgaaa
2461 ttttgtcata ccacaattgt gattcacaaa ctcgaaatgc tccagaattg gattgcctta
2521 tcagacttca ggctcagaac atacagcaac gacacgtagt ctttttaaca gagaagaaca
2581 tcaagatgct ttccagttat acagataatg gaatagtggg tgcaactgct gaagacttca
2641 tgcaaaaactt taaaaatctt gtgggctatc acaattcaat cacagaagaa aaccttccac
2701 agcttggtgc taatgagaat cttgagtcgc agtcagctct tttagaaaac gatgaaaagg
2761 atgaagagga tatgtctctg gattcagggg atgaaatctc acatatagaa gtatgcagca
2821 attttcattc agaaatatgg gagaaagaga ccaaaggatc acgtggaaca gatcaaaaaa
2881 agaatactca aattgagttg caatcgtctc ctgatgtgca aaacagttta ttagaagata
2941 agacttacct tgattctgaa gagagaactt ctattgatat agtatgctct gaaggagaga
3001 acagcaattc aacagaacaa gattcatata gtaactttca ggtttatcat agtcaattaa
3061 atatgtccca tcagtttagt cattttaatg ttctcactca tcagacattt ttggggacac
3121 catatgcctt ttcataaggt cagtctcaag aaaatgagaa ttacttctta tctgcttata
3181 ctgaaagctt ggatagagat aaatctccac ctcccttaag ttgggggaaa agtgattcct
3241 ccaggccata ttcacaagag aaataactgt agtaactttt tttttaagag attgttgtgg
3301 actttgttta ttaacaattt atatttcatt ctctaaacaa aaggttcttg ttctttctca
3361 aatgtttttt cttttattta aatcatgatg gcctgtaaca gttgaagcat ctaaaaattg
3421 aaataaatat atatttttaa cataaaaaaa aaaaaaaaaa aa

//

se89-1 protein 1035aa

1 mfpysrldd kkflysapr n kshidtlha yifrpevyql picklkelfe enrklqqfsp
61 lsdyegqeee mngtkmkfkg rnnsrgeaai sgkqrsshs dydkdrvkel inliqcrkks
121 vggdsdtedm rsktvkkrkl edlpenmrkl aktsnlsenc hlyeespppi gslghdadlr
181 rqqqdtcnsg iadihrlnfw lsetlanarh sdaslttdvn kalglstdda yeelrqkhey
241 einstpdkkd yeqptcakve naqfkgtsl llevdatsky svaistsevg tdkhlhiked
301 pnlisvnnfe dcslcpsvpi ehgfrrqgsk snnveeteih wkclipitdtl kgtteddvlt
361 gqveeqcvpa aeaepavse ttervtlgey nlfsrkieei lkqknvsyvs rvstpifstg
421 ekmkrlsefi ysktskagvq efvdglhekl ntiikasak ggnlppvspn dsgakiasnp
481 lerhvipvss sdfnnkhll plcsdplkdt nsdeqhstsa ltevemnpq hatelmvtsd
541 hivpgdmare pveettksp dvnisagpal snfisqlepe vfnslvkimk dvqkntvkfy
601 iheeeesvlc keikeylikl gntechpeqf lerrskldkl liiiqnedia gfihkipglv
661 tlkkllpcvsf agvdslddvk nhtynelfvs ggfvsvdesi lnpevvtven lknfltflee
721 lstpegkwqw kvhckfqkkl kelgrlnaka lslltllnv qkhlveils yhncdsqtrn
781 apeldclirl qaqniqqrhv vflteknim lssytdngiv vataedfmqn fknlvgyhns
841 iteenlpqlg anenlesgsa llendekdee dmsldsgdei shievcsnfh seiweketkg
901 srgtdqkknt qielqsspdv qnsllledkty ldseertsid ivcsegensn steqdsysnf
961 qvyhsqlnms hqfshfnvlt hqtflgtpya lsssqsqene nyflsaytes ldrdksppl
1021 swgksdssrp ysgek

//

Fig. 15 (2)

L14-2 2171 bp

BASE COUNT 936 a 327 c 458 g 450 t

ORIGIN

```
1 aattcggcac gaggattctt gtgccaaaac agacataggc tcagaaaatt ctgaacaaat
61 agctaatttt cctagtggaa attttgctaa acataatttca aaaacaaatg aaacagaaca
121 gaaagtaaca caaatattgg tggaattaag gtcattctaca tttccagaat cagctaataga
181 aaagacttat tcagaaagcc cctatgatac agactgcacc aagaaattta tttcaaaaat
241 aaagagcgtt tcagcatcag aggatttggt ggaagaaata gaattctgagc tcttatctac
301 ggagtttgca gaacatcaag taccaaattg aatgaataag ggagaacatg cattagttct
361 gtttgaaaag tgtgtgcaag ataaatatatt gcagcaggaa catatcataa aaaagttaat
421 taaagaaaat aagaagcatc aggagctctt cgtagacatt tgttcagaaa aagacaattt
481 aagagaagaa ctaaagaaaa gaacagaaac tgagaagcag catatgaaca caattaaaca
541 gttagaatca agaatagaag aacttaataa agaagttaaa gcttcagag ataaactaat
601 agctcaagac gttacagcta aaaatgcagt tcagcagtta caciaagaga tggcccaacg
661 gatggaacag gccaacaaga aatgtgaaga ggcacgcaa gaaaaagaag caatggtaat
721 gaaatatgta agaggtgaga aggaattctt agatcttcga aaggaaaaag agacacttga
781 gaaaaaactt agagatgcaa ataaggaact tgagaaaaac actaacaaaa ttaagcagct
841 ttctcaggag aaaggacggt tgcaccagct gtatgaaact aaggaaggcg aaacgactag
901 actcatcaga gaaatagaca aattaaagga agacattaac tctcacgtca tcaaagtaaa
961 gtgggcacaa aacaaattaa agctgaaat ggattcacac aaggaaacca aagataaact
1021 caaagaaaca acaacaaaat taacacaagc aaaggaagaa gcagatcaga tacgaaaaaa
1081 ctgtcaggat atgataaaaa catatcagga gtcagaagaa attaaatcaa atgagcttga
1141 tgcaaagctt agagtcacaa aaggagaact tgaaaaacaa atgcaagaaa aatctgacca
1201 gctagagatg catcatgcca aaataaagga actagaagat ctgaagagaa catttaagga
1261 gggatatgat gagttaagaa cactgagaac aaaggtgaaa tgtctagaag atgaacgatt
1321 aagaacagaa gatgaattat caaaatataa ggaaattatt aatcgccaaa agctgaaat
1381 tcagaattta ttggacaagg tgaaaactgc agatcagcta caggagcagc ttcaaagagg
1441 taagcaagaa attgaaaatt tgaaagaaga agtggaagt cttaattctt tgattaatga
1501 cctacaaaaa gacatcgaag gcagtaggaa aagagaatct gagctgctgc tgtttacaga
1561 aaggctcact agtaagaatg cacagcttca gtctgaatcc aattctttgc agtcacaatt
1621 tgataaagtt tcctgtagtg aaagtcagtt acaaagccag tgtgaacaaa tgaacagac
1681 aaatattaat ttggaaagta ggttggtgaa agaggaagaa ctgcgaaaag aggaagtcca
1741 aactctgcaa gctgaactcg cttgtagaca aacagaagtt aaagcattga gtaccaggt
1801 agaagaatta aaagatgagt tagtaactca gagacgtaaa catgcctcta gtatcaagg
1861 tctcaccaaa caacttcagc aagcacgaag aaaattagat cagggttgaga gtggaagcta
1921 tgacaaagaa gtcagcagca tgggaagtcg ttctagttca tcagggtccc tgaatgctcg
1981 aagcagtga gaagatcgat ctccagaaaa tactgggtcc tcagtagctg tggataactt
2041 tccacaagta gataaggcca tggtgattga gagaatagtt aggctgcaaa aagcacatgc
2101 ccggaaaaat gaaaagatag aatttatgga ggaccacatc aaacaactgg tggaaaaaaa
2161 aaaaaaaaaa a
```

//

Fig. 16 (1)

L14-2 protein 613aa

1 mnkgehalvl fekcvcqdkyl qqehiikkli kenkkhqelf vdicsekdn1 reelkkrtet
61 ekqhmntikq lesrieelnk evkasrdkli aqdvtaqnav qqlhkemaqr meqankkcee
121 arqekeamvm kyvrgekesl dlrkeketle kklrdankel ekntnkikql sqekgrlhql
181 yetkegettr lireidklke dinshvikvk waqnklkaem dshketkdki ketttkltqa
241 keeadqirkn cqdmiktyqe seeiksneld aklrvtkgel ekqmgekstdq lemhhakike
301 ledlkrtfke gmdelrtlrk kvkclederi rtedelskyk eiinrqkaei qnlldkvkta
361 dqlqeqlqrg kqeienlkee veslnslind lqkdiegsrk reselllfte rltsknaqlq
421 sesnslqsqf dkvscsesql qsqceqmktq ninlesrllk eeelrkeevq tlqaelacrq
481 tevkaltstqv eelkdelvtq rrrkassikd ltkqlqqarr kldqvesgsy dkevssmgsr
541 ssssgslnar ssaedrpen tgssvavdnf pqvdkamlie rivrlqkaha rknekiefme
601 dhikqlvekk kkk

//

Fig. 16 (2)

L15-7 3513 bp

BASE COUNT 1234 a 723 c 867 g 689 t

ORIGIN

```
1  ctcgaaatta accctcacta aagggaaaca aagctggagc tccaccgagg tggcgccgcg
61  tctagaacta gtggatcccc cgggctgcag gaattcggca cgaggatgat ctgctgctgc
121 tgctgctgct gccgcccgcg cctctattgc tgatactcta gtggggctgg aagggtggtt
181 cctattcgca ccatcgccaa ccagagacag agggaaaaaa aaaaccggca gccactgctg
241 atgttgggtt cggaggctgc atccgactcg gtcacaagga aaatggattc agtttgcac
301 tctcctcct ttaaacagct tctccgggtc tcagcatggg cttccagggc agcgattgag
361 gagaccttac caaggagcac cacacagtag atgctgagac atcgtactcc aggataagaa
421 acagtaacat ggcagcacct gcttgaaaga aattaaaaac caacagactc catttagaaa
481 ggaacaatgt ccaagaaagg gcgaaataag ggcgagaagc ccgaggcact cattgttgcc
541 cttcaagctg ccaatgaaga cctcaggacc aagctcacag acattcagat agagctgcat
601 caagagaagt ccaaggtatc aaagcttgaa agagagaaga ctcaagaagc gaagaggatt
661 cgtgagctgg agcagcgcaa gcacacggtg ctggtgacag aactcaaagc caagctccat
721 gaggagaaga tgaaggagct gcaggctgtg agggagaacc ttatcaagca gcacgagcag
781 gaaatgtcaa ggacggtgaa ggtacgtgat ggagagatcc agaggctcaa gtctgctctc
841 tgtgctctcc gcgacggcag cagtgacaaa gtaaggacag cgctcaccat tgaggcccgg
901 gaggaggccc ggaaactgtt tgacacagag cgccttaagc tcttacagga aattgcggac
961 ctgaaaacgg ccaagaagca ggtggacgag gctctgagca atatgatcca agcagataaa
1021 atcaaggctg gggaccttcg gagtgagcat cagtcaccac aagaagccat ctogaagatc
1081 aagtgggagt cggagcggga tattcgaggc ctgatggatg aaatcaaagc caaggacagg
1141 atcatctttt ccctggaaaa ggaactggag acccagacag gctatgtaca gaaactccaa
1201 cttcagaagg aggccttgga cgaacaactc tttctggtca aggaggctga gtgcaacatg
1261 agcagcccaa aacgagaaat tccaggaagg gcaggatgat gttccgaaca ctgcagcagt
1321 cctgatttgc gaagaaatca aaagagaata gctgaattga atgccactat aagaaaatta
1381 gaagacagga ataccttgct tggagatgaa cgaaatgaac tggtaaaacg tgtgcgggaa
1441 accgaaaagc aatgtaaacc tctcctggaa aggaacaagt gcctcgccaa gagaaacgat
1501 gaactgatgg tgtccttgca gcgcatggaa gaaaaactaa aagccgttac caaggaaaat
1561 tcagaaatga gagaaaaaat aacatcccat ccaccctga agaaattaaa atctctgaat
1621 gacctcgacc aagctaataga agaacaagaa acagagtttc taaaacttca ggtcattgag
1681 caacagaaca ttattgatga gctcacaagg gaccgagaaa agctcatccg tagaagaaag
1741 catagaagaa gttccaagcc aattaagagg cctgttttgg acccgtttat tggctatgat
1801 gaggactcta tggattcaga gacatcatcc atggcctcat ttagaacaga cagaacacca
1861 gctactcctg atgatgactt ggatgaaagt ttagcagctg aagaatctga actaagattt
1921 cgacaattaa caaaagaata tcaggccctc caaagagcat atgccctcct acaggagcag
1981 acgggaggca tcatcgacgc tgaacgagaa gccaaaggctc aagaacagct ccaagcagag
2041 gtgctaaggt ataaagccaa aattgaagac ctggaagcga ctctggctca gaaagggcag
2101 gattcacact gggtagaaga taaacaactt ttcattaaga gaaaccagga gcttttagaa
2161 aagatagaaa aacaggaggc agaaaatcac cggttacaac aagaactaca ggacgccaga
2221 gaccagaatg agctgctgga gtttcgaaac ctagagctag aagagagaga gagacgatcc
```

This Page Blank (uspto)

2281 cctccattta atctccaaat tcaccattc tcagatggtg tgagtgtctt acagatctac
2341 tgtatgaaag aaggtgttaa ggatgtgaac atccctgac tcataaagca gcttgatata
2401 ttgggtgata atgggaattt aagaaatgaa gaacaagtgg ccataattca ggccagcact
2461 gtgctgtccc tggcagagaa gtggatccag cagattgaag gagctgaggc tgccctacac
2521 cagaaaatga tggaattgga aagtgcacatg gaacagttct gcaaaataaa aggctatctg
2581 gaggaagaac tagactacag aaaacaagct cttgaccaag catatatgag aatccaggaa
2641 ctagaagcta ctttgtacaa tgctctacag caagaaactg ttatcaagtt tgggtgaatta
2701 ttaagtgaag aacagcaaga ggagctgagg acggcagtag aaaagttacg gcggcaaagt
2761 ctgaggaaga gcagagagta tgactgtcag attcttcagg agagaatgga gctcttacag
2821 caagcccatc agagaattcg tgacttagaa gataaaacag acatccagaa aagacaaata
2881 aaagacttag aagaaaagag taaccgaaaa catggataag atcccaggaa gacaagtgtc
2941 tctaaacctt caaagatggc aaaattgttt acaccagtga gagggagatc aaaagctaag
3001 aactaccctg tagccaggac tacaactgtg tatttttaaag ccattattca aggtttctta
3061 cttgacagtt cctacacaac cctgttgaaa atctacaata tatgtgtcat ttaatgaaac
3121 atgtatatgt caaatcagaa gagaagaact ataaacatat attgtgtaaa gaaaaagttc
3181 agcaatggaa ctagtctctg cagatcaagc aaagatgtgt cttgggcatg gaaccaaagt
3241 tacaatgaaa tattcaacct ctgctgtgca ggggggtcat tttaatgtaa caccacacct
3301 catggaaaca ctagtctctg taataaacat catttttaaaa gatcaaaaca aacaaacaaa
3361 aaaaacaagg gtgggtgggg agtgaagcac gaggaatacc tatgaagagc tatttacaat
3421 aaaatgtttc atttgaaaag tcaaaaaaaaa aaaaaaaaaa actcgagggg gggcccggtg
3481 cccaattcgc cctatagtga gtcgtattac aat

//

L15-7 protein 810aa

1 mskkgrnkge kpealivalq aanedlrkl tdiqielhqe kskvsklere ktqeakrire
61 leqrkhtvlv telkaklhee kmkelqavre nliqhegem srtvkvrde iqrlksalca
121 lrdgssdkvr taltiearee arklfdterl kllqeiadlk takkqvdeal snmiqadkik
181 agdlrsehqs hgeaaiskikw eserdirrlm deikakdrii fslekeletq tgyvqklqlq
241 kealdeqlfl vkeaecnms pkreipgrag dgsehcsspd lrrnqkriae lnatirkled
301 rntllgdern ellkrvrete kqckpllern kclakrnde mvslqrmeek lkavtkense
361 mrekitshpp lklklsndl dqaneeqete flklqviegq niideltrdr eklirrrkhr
421 rsskpikrpv ldpfigyded smdsetssma sfrtdrtpat pddldesla aeeselrfrq
481 ltkeyqalqr ayallqeqtg giidaereak ageqlqaevl rykakiedle atlaqkgqds
541 hwvedkqlfi krnqelleki ekqeaenhrl qqelqdardq nellefrnle leererrssp
601 fnlqihpfsd gvsalqiycm kegkvdnip dlikqldilg dngnlrnee vaiiqastvl
661 slaekwiqqi egaeaalhqk mmelesdmeq fckikgylee eldyrkqald qaymriqele
721 atlynalqge tvikfgells ekqgeelrta veklrrqmlr ksreydcqil qermellqqa
781 hqirdledk tdiqkrqikd leeksnrkhg

//

Fig. 17 (2)

Li9-1 1465 bp

BASE COUNT 334 a 423 c 443 g 265 t

ORIGIN

```
1 aattcggcac gaggcgcggg ccgctgtgag ggcgcggcggc gagcgacggg cgcggggccg
61 cggagcagcg agcgagcag cgagcgcgag gccggagccc cggccaggcc cggccgaccc
121 gccgagcccc cgatgcgccc cggggccgcc ccccggcgca gctgacgcc cgcggccccg
181 cgaagacccc ggccggcccg tcccggagga agcggccgcc gccgcgcgcg cccagcccg
241 cgcgcgcgcc gcccgggcac catggcgggg aaggcggccg ccccgggcac cgcggtgctg
301 ctggtcacgg ccaacgtggg ctgctcttc gacgacccag aaaacctgca gaagaactgg
361 cttcgggaat tttaccaggt cgtgcacaca cacaggccgc acttcatggc cttgcaactg
421 caggagtttg gaggaagaa ctacgaggcc tccatgtccc acgtggacaa gttcgtcaaa
481 gaactattgt cgagtgatgc gatgaaagaa tataacaggg ctcgagtcta cctggatgaa
541 aactacaaat cccaggagca cttcacggca ctaggaagct tttattttct tcatgagtc
601 ttaaaaaaca tctaccagtt tgactttaaa gctaagaagt atagaaaggc cgctggcaaa
661 gagatctact cggatacctt agagagcacg cccatgctgg agaaggagaa gtttcccgag
721 gactacttcc ccgagtgcaa atgggtcaaga aaaggcttca tccggacgag gtggtgcatt
781 gcagactgtg cctttgactt ggtgaatata catcttttcc atgatgcttc caatctggtc
841 gcctgggaaa caagcccttc cgtgtactcg ggaatccggc acaaggcact gggctacgtg
901 ctggacagaa tcattgatca gcgattcgag aaggtttctt actttgtatt tgggtatttc
961 aacttccggc tggattccaa gtccgtcgtg gagacgctct gcacaaaagc caccatgcag
1021 acggtccggg ccgccgacac caatgaagtg gtgaagctca tatttcgtga gtcggacaac
1081 gaccggaagg ttatgctcca gttagaaaag aaactcttcg actacttcaa ccaggagggt
1141 ttccgagaca acaacggcac cgcgctcttg gagtttgaca aggagttgtc tgtctttaag
1201 gacagactgt atgaactgga catctcgttc cctcccagct acccgtagag tgaggacgcc
1261 cgccagggtg agcagtacat gaacaccggg tgcccagcct ggtgtgaccg catcctcatg
1321 tccccgtctg ccaaggagct ggtgctgcgg gtgagtggtg gctgccccag ccctgggcac
1381 agaggggatg ggagcgtgg gtctggtctg gccagccct ggtgacaggg cccaggggt
1441 gggggaaaaa aaaaaaaaaa aaaaa
```

//

L9-1 protein 387aa

```
1 magkaaapgt avllvtanvg slfddpenlq knwlrefyqv vthhrphfma lhcqefggkn
61 yeasmshvdk fvkellssda mkeynrarvy ldenyksqeh ftalgsfyfl heslkniyqf
121 dfkakkyrkv agkeiysdtl estpmlekek fpqdyfpeck wsrkgfirtr wciadcafdl
181 vnihlfhdas nlvawetsps vysgirhkal gyvldriidq rfekvsyfvf gdfnfrldsk
241 svvetlctka tmqtvraadn nevvklifre sdndrkvmllg lekklfdyfn qevfrdnngt
301 allefdkels vfkdrlyeld isfppsypys edarqgeqym ntrcpawcdr ilmspsakel
361 vlrvsvccps pghrgmwsag sglaqpw
```

//

Fig. 18

Li9-4 1681bp

BASE COUNT 545 a 304 c 327 g 505 t

ORIGIN

```
1 gaggggttaga tcgagcaacc ctctaaaagc agtttagagt ggtaaaaaaa aaaaaaacac
61 accaaacgct cgcagccaca aaagggatga aatttcttct ggacatcctc ctgcttctcc
121 cgttactgat cgtctgctcc ctagagtcct tcgtgaagct ttttattcct aagaggagaa
181 aatcagtcac cggcgaaatc gtgctgatta caggagctgg gcatggaatt gggagactga
241 ctgcctatga atttgctaaa cttaaaagca agctggttct ctgggatata aataagcatg
301 gactggagga aacagctgcc aaatgcaagg gactgggtgc caaggttcat acctttgtgg
361 tagactgcag caaccgagaa gatatttaca gctctgcaaa gaaggtgaag gcagaaattg
421 gagatgtagt tatttttagta aataatgctg gtgtagtcta tacatcagat ttgtttgcta
481 cacaagatcc tcagattgaa aagacttttg aagttaatgt acttgcacat ttctggacta
541 caaaggcatt tcttcctgca atgacgaaga ataaccatgg ccatattgtc actgtggctt
601 cggcagctgg acatgtttcg gtcccttct tactggctta ctgttcaagc aagtttgctg
661 ctgttggttgc tcataaaact ttgacagatg aactggctgc cttacaaata actggagtca
721 aaacaacatg tctgtgtcct aatttcgtaa aactggctt catcaaaaat ccaagtacaa
781 gtttgggacc cactctggaa cctgaggaag tggtaaacag gctgatgcat gggattctga
841 ctgagcagaa gatgattttt attccatctt ctatagcttt tttaacaaca ttggaaagga
901 tccttcctga gcgtttcctg gcagttttaa aacgaaaaat cagtgttaag tttgatgcag
961 ttattggata taaaatgaaa gcgcaataag cacctagttt tctgaaaact gatttaccag
1021 gtttaggttg atgtcatcta atagtgccag aattttaatg tttgaacttc tgttttttct
1081 aattatcccc atttcttcaa tatcattttt gaggcttttg cagtcttcat ttactaccac
1141 ttgttcttta gccaaaagct gattacatat gatataaaca gagaaatacc tttagagggtg
1201 actttaagga aatgaagaa aaagaaccaa aatgacttta ttaaataat ttccaagatt
1261 atttgtggct cacctgaagg ctttgcaaaa tttgtaccat aaccgtttat ttaacatata
1321 tttttatttt tgattgcact taaattttgt ataattttgt tttctttttc tgttctacat
1381 aaaatcagaa acttcaagct ctctaaataa aatgaaggac tatatctagt ggtatttcac
1441 aatgaatata atgaactctc aatgggtagg ttcatccta ccattgcca ctctgtttcc
1501 tgagagatac ctcacattcc aatgccaaac atttctgcac agggaagcta gaggtggata
1561 cacgtgttgc aagtataaaa gcatcactgg gatttaagga gaattgagag aatgtaccca
1621 caaatggcag caataataaa tggatcacac ttaaaaaaaaa aaaaaaaaaa aaaaaaaaaa
1681 a
```

//

Li9-4 protein 300aa

```
1 mkfllldilll lpllivcsle sfvklfipkr rksvtgeivl itgaghgigr ltayefaklk
61 sklvldink hgleetaakc kglgakvhtf vvdcsnredi yssakkvkae igdvsilvnn
121 agvvytsdlf atqdpqiekt fevnvlahfw ttkafllpamt knnhghivtv asaaghvsvp
181 fllyacsskf aavgfhkltt delaalqitg vkttcclcpnf vntgfiknps tslgptlepe
241 evvnrlmhgi lteqkmifip ssiaflttle rilperflav lkrkiskvkd avigykmaq
301
```

//

Fig. 19

38/44

Lii5-2

BASE COUNT 341 a 350 c 362 g 315 t

ORIGIN

```
1 aattcggcac gaggggagcg cagcagccat ggcaagccgt ctctgctca acaacggcgc
61 caagatgccc atcctggggt tgggtacctg gaagtcacct ccagggcagg tgactgaggg
121 cgtgaagggtg gccattgacg tgggtaccg ccacatcgac tgtgcccattg tgtaccagaa
181 tgagaatgag gtgggggtgg ccattcagga gaagctcagg gagcagggtgg tgaagcgtga
241 ggagctcttc atcgtcagca agctgtggtg cacgtacct gaggaaggcc tggtgaaagg
301 agcctgccag aagacactca gcgacctgaa gctggactac ctggacctct acctattca
361 ctggccgact ggctttaagc ctgggaagga atttttccca ttggatgagt cgggcaatgt
421 ggttcccagt gacaccaaca ttctggacac gtgggcgggc atggaagagc tggtggaatga
481 agggctggtg aaagctattg gcatctccaa cttcaaccat ctccagggtg agatgatctt
541 aaacaaacct ggcttgaagt ataagcctgc agttaaccag attgagtgcc acccatatct
601 cactcaggag aagttaatcc agtactgcc gtccaaaggc atcgtggtga ccgcctacag
661 cccctcggc tctcctgaca ggccctgggc caagcccgag gacccttctc tcttgaggga
721 tcccaggatc aaggcgatcg cagccaagca caataaaact acagcccagg tctgatccg
781 gttcccatg cagaggaact tgggtggtgat cccaagtct gtgacaccag aacgcattgc
841 tgagaacttt aaggctcttg actttgaact gagcagccag gatatgacca cttactcag
901 ctacaacagg aactggaggg tctgtgcctt gttgagctgt acctcccaca aggattacct
961 cttccatgaa gagttttgaa gctgtggtt cctgctcgtc cccaagtgc ctatacctgt
1021 gtttcttgcc tcatTTTTT ccttgcaaat gtagtatggc ctgtgtcact cagcagtggg
1081 acagcaacct gtagagtggc cagcgagggc gtgtctagct tgatgttgga tctcaagagc
1141 cctgtcagta gagtagaagt ctcttccagt ttgctttgcc cttctttcta cctgctggg
1201 gaaagtacaa cctgaatacc cttttctgac caaagagaag caaatctac caggtcaaaa
1261 tagtgccact aacggttgag ttttgactgc ttggaactgg aatcctttca gcaagacttc
1321 tctttgcctc aaataaaaag tgcttttgtg aaaaaaaaaa aaaaaaaa
```

//

Lii5-2 protein 316aa

```
1 masrlllnng akmpilglgt wksppgqvte avkvaidvgy rhidcahvyg nenevgvaiq
61 eklreqvvkr eelfivsklw ctyhekglvk gacqktlsdl kldyldlyli hwptgfkpgk
121 effpldesgn vvpstnild twaameelvd eglvkaigis nfnhlqvemi lnkpglkykp
181 avngiechpy ltqekliqyc qskgivvtay splgspdrpw akpedpslle dprikaiaak
241 hnkttaqvli ifpmqrnlv ipksvtperi aenfkvfdf lssqdmtil synrnwrvca
301 llstshkdy pfheef
```

//

Fig. 20

39/44

Li110-6

BASE COUNT . 431 a 308 c 364 g 391 t

ORIGIN

```
1 aattcggcac gaggggtggaa atttgagacc agcaagtact atgtgactat cattgatgcc
61 ccaggacaca gagactttat caaaaacatg attacagggg catctcaggg tgactgtgct
121 gtcctgattg ttgctgctgg tgttggtgaa tttgaagctg gtatctccaa gaatggggcag
181 acccgagagc atgcccttct ggcttacaca ctgggtgtga aacaactaat tgtcgggtgtt
241 aacaaaatgg attccactga gccaccctac agccagaaga gatatgagga aattgttaag
301 gaagtcagca cttacattaa gaaaattggc tacaaccccg acacagtagc atttgtgcca
361 atttctgggtt ggaatggtga caacatgctg gagccaagtg ctaacatgcc ttgggtcaag
421 ggatggaaag tcacccgtaa ggatggcaat gccagtggaa ccacgctgct tgaggctctg
481 gactgcatcc taccaccaac tcgtccaact gacaagccct tgcgcctgcc tctccaggat
541 gtctacaaaa ttggtggtat tgggtactgtt cctgttggcc gagtgagagc tgggtgttctc
601 aaacccggta tgggtggtcac ctttgcctca gtcaacgta caacggaagt aaaatctgtc
661 gaaatgcacc atgaagcttt gagtgaagct cttcctgggg acaatgtggg cttcaatgtc
721 aagaatgtgt ctgtcaagga tgttcgtcgt ggcaacgttg ctggtgacag caaaaatgac
781 ccaccaatgg aagcagctgg cttcactgct caggtgatta tcctgaacca tccaggccaa
841 ataagcgccg gctatgcccc tgtattggat tgccacacgg ctcacattgc atgcaagttt
901 gctgagctga aggaaaagat tgatcgccgt tctggtaaaa agctggaaga tggccctaaa
961 ttcttgaagt ctggtgatgc tgccattggt gatatgggtc ctggcaagcc catgtgtgtt
1021 gagagcttct cagactatcc acctttgggt cgctttgctg ttcgtgatat gagacagaca
1081 gttgcggtgg gtgtcatcaa agcagtggac aagaaggctg ctggagctgg caaggtcacc
1141 aagtctgccc agaaagctca gaaggctaaa tgaatattat ccctaatacc tgccacccca
1201 ctcttaataca gtggtggaag aacggtctca gaactgtttg tttcaattgg ccattttaagt
1261 ttagtagtaa aagactggtt aatgataaca atgcatcgta aaaccttcag aaggaaagga
1321 gaatgttttg tggaccactt tggttttctt ttttgctgtg ggcagtttta agttattagt
1381 ttttaaaatc agtacttttt aatggaaaca acttgaccaa aaatttgtca cagaattttg
1441 agaccatta aaaaagttaa atgagaaaaa aaaaaaaaaa aaaaaaaaaa aaaa
```

//

Li110-6 protein 361aa

```
1 mitgtsqadc avlivaagvg efeagiskng qtrehallay tlgvkqlivg vnkmdstepp
61 ysqkryeeiv kevstyikki gypndtvafv pisgwnngdm lepsanmpwf kgwkvtrkdg
121 nasgttlllea ldcilpptrp tdkplrlplq dvykiggigt vpvgrvetgv lkpgmvvtfa
181 pvnvttevks vemhhealse alpgdnvgfn vknvsvkdvr rgnvagdskn dppmeaagft
241 aqviilnhpg qisagyapvl dchtahiack faelkekidr rsgkkledgp kflksгдаai
301 vdmvpgkpmc vesfsdypl grfavrdmrq tvavgvikav dkkaagagkv tksaqkaqka
361 k
```

//

Fig. 21

40/44

Liii4-5

BASE COUNT 113 a 84 c 99 g 94 t 8 others

ORIGIN

```
1  ngthctbrty nchckaattc ggcacgaggc tgcggacata aatcttaaag ctagtaacat
61 gttgttcttc taggaattcc attcagctac agatttaagg tttatcagta gtatttccag
121 aaagatggtc cgacacagtg gctcacgttt ataatcccag cactttggga ggccgaggtg
181 ggtgaattgc ttgagtccag gagttcaaga ccagcctggg caacatggca aaaccctgtc
241 tttgcctgta gtacccccag ctatttgaga ggctgaggtg gaagaatcac ctgagcctgg
301 ggaggtcagg gctgcagtgt gctgaaattg cacaactgca ctccagcctg ggcaatcaga
361 gtgagaccct gtctttaaga aaaaaaaaaa aaaaaaaaaa
```

//

Fig. 22

Nukleotid-Sequenz von GBP-TA (die Splicing Variante GBP-TA_{short} hat eine Lücke von Nukleotid Nummer 411 bis einschließlich 518):

```
1  GCACTGAGGT CACCCTCCAG GCTGTGGAAC CTTTGTCTTT TCACTCTTTG
51  CAATAAATCT TGCTGCTGCT CACTCTTTGG GTCCACACTG CCTTTATGAG
101 CTGTAACACT CACTGGGAAT GTCTGCAGCT TCACTCCTGA AGCCAGCGAG
151 ACCACGAACC CACCAGGAGG AACAAACAAC TCCAGACGCG CAGCCTTAAG
201 AGCTGTAACA CTCACCGCGA AGGTCTGCAG CTTCACTCCT GAGCCAGCCA
251 GACCACGAAC CCACCAGAAG GAAGAACTC CAAACACATC CGAACATCAG
301 AAGGAGCAAA CTCCTGACAC GCCACCTTTA AGAACCGTGA CACTCAACGC
351 TAGGGTCCGC GGCTTCATTC TTGAAGTCAG TGAGACCAAG AACCCACCAA
401 TTCCGGACAC GCTAATTGTT GTAGATCATC ACTTCAAGGT GCCCATATCT
451 TTCTAGTGGA AAAATTATTC TGGCCTCCGC TGCATACAAA TCAGGCAACC
501 AGAATTCTAC ATATATAAGG CAAAGTAACA TCCTAGACAT GGCTTTAGAG
551 ATCCACATGT CAGACCCCAT GTGCCTCATC GAGAACTTTA ATGAGCAGCT
601 GAAGGTTAAT CAGGAAGCTT TGGAGATCCT GTCTGCCATT ACGCAACCTG
651 TAGTTGTGGT AGCGATTGTG GGCCTCTATC GCACTGGCAA ATCCTACCTG
701 ATGAACAAGC TGGCTGGGAA GAACAAGGGC TTCTCTGTTG CATCTACGGT
751 GCAGTCTCAC ACCAAGGGAA TTTGGATATG GTGTGTGCCT CATCCCAACT
801 GGCCAAATCA CACATTAGTT CTGCTTGACA CCGAGGGCCT GGGAGATGTA
851 GAGAAGGCTG ACAACAAGAA TGATATCCAG ATCTTTGCAC TGGCACTCTT
901 ACTGAGCAGC ACCTTTGTGT ACAATACTGT GAACAAAATT GATCAGGGTG
951 CTATCGACCT ACTGCACAAT GTGACAGAAC TGACAGATCT GCTCAAGGCA
1001 AGAAACTCAC CCGACCTTGA CAGGGTTGAA GATCCTGCTG ACTCTGCGAG
1051 CTTCTTCCCA GACTTAGTGT GGACTCTGAG AGATTTCTGC TTAGGCCTGG
1101 AAATAGATGG GCAACTTGTC ACACCAGATG AATACCTGGA GAATTCCCTA
1151 AGGCCAAAGC AAGGTAGTGA TCAAAGAGTT CAAAATTTCA ATTTGCCCCG
1201 TCTGTGTATA CAGAAGTTCT TTCCAAAAAA GAAATGCTTT ATCTTTGACT
1251 TACCTGCTCA CCAAAAAAAG CTTGCCCAAC TTGAAACACT GCCTGATGAT
1301 GAGCTAGAGC CTGAATTTGT GCAACAAGTG ACAGAATTCT GTTCCTACAT
1351 CTTTAGCCAT TCTATGACCA AGACTCTTCC AGGTGGCATC ATGGTCAATG
```

Fig. 23

Abgeleitete Proteinssequenz von GBP-TA:

1401 GATCTCGTCT AAAGAACCCTG GTGCTGACCT ATGTCAATGC CATCAGCAGT
 1451 GGGGATCTGC CTTGCATAGA GAATGCAGTC CTGGCCCTTG CTCAAGAGA
 1501 GAAGTCAAGT GAGTGCAGG AGGCCATTGC CCACTATGAC CAGCAATGG
 1551 GCCAGAAAGT GCAGCTGCC ATGGAACCC TCAGGAGCT GCTGGACCTG
 1601 CACAGGACCA GTGAGAGGGA GCCATTGAA GTCTTCATGA AAAACTCTTT
 1651 CAAGGATGTA GACCAAGTT TCAGAAAGT ATTGAGACT CTACTAGATG
 1701 CAACAAGAA TGACATTTGT AACGGAAAC TCAGAGCATC CTCGGATTAT
 1751 TGCTCGGCTT TACTTAAGGA TATTTTGGT CCTCTAGAAG AAGCAGTGAA
 1801 GCAGGAATT TATTCTAAGC CAGGAGGCCA TAATCTCTTC ATTCAGAAAA
 1851 CAGAAAGACT GAAGGCAAG TACTATCGGG AGCCTCGGAA AGGAATACAG
 1901 GCTGAAGAAG TTCTGCAAGAA ATATTTAAG TCAGAAGAGT CTGTGAGTCA
 1951 TGCATATTTA CAGACTGACC AGGCTCTCAC AGAGACGGAA AAAAGAGAAG
 2001 AAGAGGACCA AGTGAAGCA GAAGCTGAAA AGGCTGAAAG GCAAAAGTTG
 2051 GCGGCGATTG AAAGGCAAG CAGCAGAAATG ATGCAAGAGA GGAAGAGACT
 2101 CCATCAGGAA CAGGTGAGAC AAATGAGAT AGCCAAACAA AATTGGCTGG
 2151 CAGAGCAACA GAAATGCAAG GAACAACAGA TGCAAGACA GGCTGCACAG
 2201 CTCAGCACAA CATGCCAAGC TCAAAATAGA AGCCTTCTCA GTAGCTCCA
 2251 GCAGGCCCAG AGGACTGTTA ATACGATGA TCATGTGTT TTAAGCTAAA
 2301 GTGCTAATA TGGAAGTTT CTTTTTTAC TCTTTGTAC TGAAGACACA
 2351 ACAGAAAGA AACTGTAGAC CTGGGACAA TCACATTTA AATAACTTT
 2401 ATATTTATTT TTCAAACTT TCAAAAA

Fig. 23 (2)

1 MATEIHMSDP MCLIHNFNEQ LKVNQEALEI LSAITQPVV VAIVGLXYTG
 51 KSYIMNKLAG KKGFSVAST VQSHTKGIWI WCVPHBPNWP HTLVLLDTLG
 101 LGDVEKADNK NDIQIFALAL LLSSTFVYNT VNKIDGALD LTHNVTELTG
 151 LTKARNSPDL DRVEBPADSA SFEPDLVWTL RPFCLGLEID GQLVTPDEYL
 201 ENSLRPKQGS DQKVQNFNL RLCIOKFPPK KKCFTFLPA HOKKLAQLET
 251 LPDDELPEEF VQVTEFCSY IFSSHMTKTL PGGIMVNGSR LKNLVLTYYN
 301 AISSGDLPCI ENAVLALAQE ENSAAVQKAI AHYDQMGOK VQLPMETLQE
 351 LLDLHRTSER EALIEVFMKNS FKVDQSFQK ELFTLLDAKQ NDICKRNLEA

43/44

401 SSDYCSALLK DIFGPLEEAV KQGIYSKPGG HNLFIQKTEE LKAKYYREPR
451 KGIQAEEVLQ KYLKSKEVS HAILQTDQAL TETEKKKKEA QVKAEEAEAE
501 AQRLAAIQRO NEQMMQERER LHQEQVRQME IAKQNWLAEO QKMQEQQMQE
551 QAAQLSTTFQ AQNRSLLEL QHAQRTVNND DPCVLL

Fig. 23 (3)

44/44

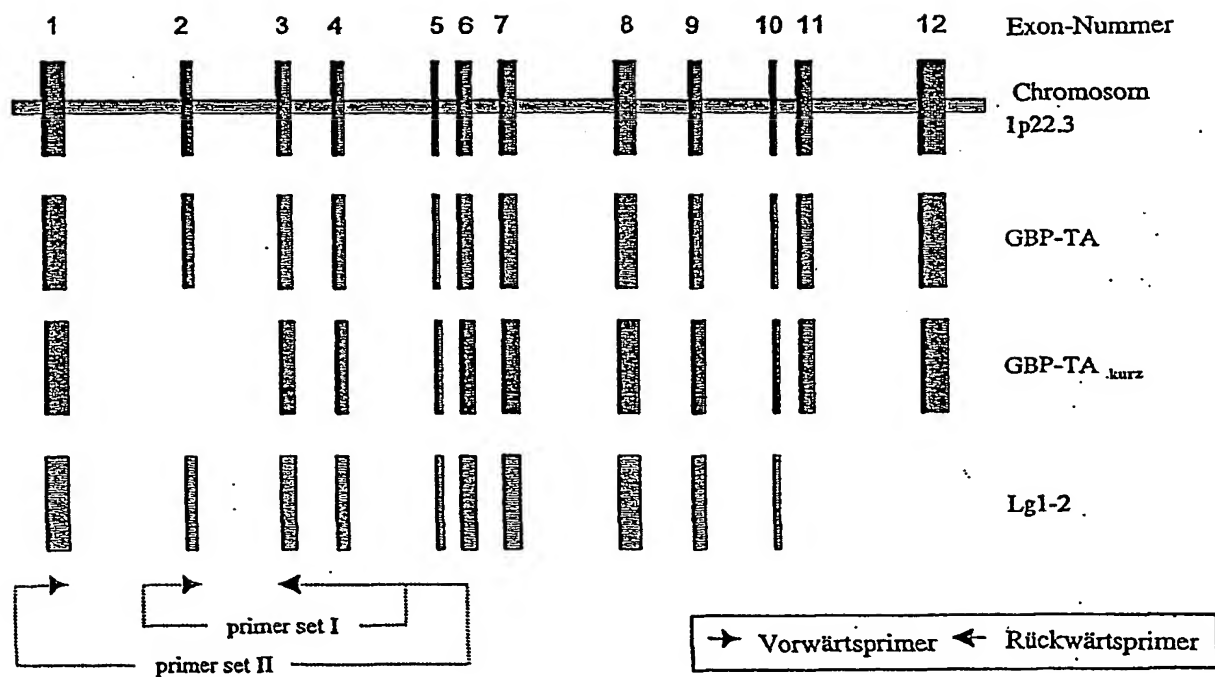


Fig. 24

**This Page is Inserted by IFW Indexing and Scanning
Operations and is not part of the Official Record**

BEST AVAILABLE IMAGES

Defective images within this document are accurate representations of the original documents submitted by the applicant.

Defects in the images include but are not limited to the items checked:

- ☒ BLACK BORDERS
- ☒ IMAGE CUT OFF AT TOP, BOTTOM OR SIDES
- ☐ FADED TEXT OR DRAWING
- ☐ BLURRED OR ILLEGIBLE TEXT OR DRAWING
- ☐ SKEWED/SLANTED IMAGES
- ☐ COLOR OR BLACK AND WHITE PHOTOGRAPHS
- ☒ GRAY SCALE DOCUMENTS
- ☒ LINES OR MARKS ON ORIGINAL DOCUMENT
- ☒ REFERENCE(S) OR EXHIBIT(S) SUBMITTED ARE POOR QUALITY
- ☐ OTHER: _____

IMAGES ARE BEST AVAILABLE COPY.

As rescanning these documents will not correct the image problems checked, please do not report these problems to the IFW Image Problem Mailbox.

This Page Blank (uspto)